

# THÈSE

Pour obtenir le grade de  
Docteur

Délivrée par le  
**Centre international d'études supérieures en sciences  
agronomiques**  
Montpellier

Préparée au sein de l'école doctorale Systèmes Intégrés en Biologie,  
Agronomie, Géosciences

Et de l'unité de recherche AGAP

Spécialité : Biologie intégrative de la plante

**Présentée par Jérôme BARTHOLOMÉ**

Déterminisme génétique de la dynamique  
de croissance et de la composition isotopique  
du carbone chez l'*Eucalyptus* en réponse aux  
variations environnementales

Soutenue le 28 mars 2014 devant le jury composé de

Catherine BASTIEN, DR, INRA Orléans	Rapporteur
René VAILLANCOURT, PR, UTAS	Rapporteur
Thierry AMEGLIO, DR, INRA Clermont-Ferrand	Examineur
Matthieu FALQUE, IR, INRA Versailles-Grignon	Examineur
Patrice THIS, DR, INRA Montpellier	Président
Jean-Marc BOUVET, Chercheur, CIRAD Montpellier	Directeur
Jean-Marc GION, Chercheur, CIRAD Bordeaux	Co-Directeur



# Résumé

Déterminisme génétique de la dynamique de croissance et de la composition isotopique du carbone chez l'*Eucalyptus* en réponse aux variations environnementales.

Les différents scénarios sur l'évolution du climat prévoient une augmentation de la fréquence et de l'intensité des sécheresses. La croissance des arbres forestiers étant fortement conditionnée par la disponibilité en eau, ces changements devraient impacter de manière significative la productivité des forêts plantées. La compréhension de l'impact des facteurs génétiques et environnementaux sur la dynamique de croissance est donc un enjeu majeur pour assurer les niveaux de production des plantations de demain. L'*Eucalyptus*, grâce à sa croissance rapide et à la disponibilité de ressources génétiques et génomiques, est un modèle biologique idéal pour mener ces recherches.

L'objectif de cette thèse est de caractériser l'architecture génétique de la dynamique de croissance à différentes échelles de temps chez l'eucalyptus en relation avec : (i) les variations environnementales, et notamment l'évolution de la disponibilité en eau, et (ii) la composition isotopique du carbone de l'arbre ( $\delta^{13}\text{C}$ ), un caractère lié à l'efficacité d'utilisation de l'eau. Pour répondre à cet objectif, un croisement interspécifique *Eucalyptus urophylla* x *E. grandis* a été étudié dans quatre dispositifs expérimentaux en République du Congo. Notre approche se base sur la cartographie des loci à effet quantitatif (QTL) et combine : (i) un génotypage haut débit, (ii) une caractérisation inter et intra-annuelle de la croissance et du  $\delta^{13}\text{C}$ , ainsi qu'un suivi en continu des micro-variations du rayon et (iii) une caractérisation en continu des facteurs environnementaux.

Ces travaux ont tout d'abord conduit à la construction des premières cartes génétiques à haute résolution chez l'*Eucalyptus*. L'analyse de l'architecture génétique du  $\delta^{13}\text{C}$  a ensuite permis de mettre en évidence des gènes candidats positionnels, potentiellement impliqués dans la variation de ce caractère. Enfin, la caractérisation inter et intra-annuelle de la dynamique de croissance a permis de montrer que l'architecture génétique de la croissance, au stade adulte, est structurée par les réponses à l'environnement au stade juvénile. Ces réponses ont ensuite été analysées grâce aux profils de micro-variations du rayon, permettant ainsi de préciser leurs déterminants génétiques.

Nos résultats soulignent l'importance de considérer la croissance comme un caractère dynamique, non seulement pour la compréhension de ses bases génétiques, mais également à des fins de sélection de variétés adaptées à un environnement changeant.

**Mots clés :** dynamique de croissance, micro-dendromètre,  $\delta^{13}\text{C}$ , cartographie génétique, QTL, interaction génotype × environnement, *Eucalyptus*.

# Abstract

Genetic determinism of growth dynamics and carbon isotope composition in *Eucalyptus* in response to environmental changes.

Scenarios of climate changes forecast an increase in frequency and intensity of droughts, related to an increase of global temperatures and changes in rainfall distribution. Growth of forest trees highly depends on water availability and will be significantly impacted by these changes. The understanding of the impact of genetic and environmental factors on the growth dynamics is a major challenge to ensure production levels of future planted forests. *Eucalyptus*, thanks to its rapid growth and the availability of genetic and genomic resources, is a perfect model to conduct this research.

The objective of this thesis is to characterize the genetic architecture of growth dynamics in *Eucalyptus* at different time scales, in relation with: (i) environmental changes, including changes in water availability, and (ii) isotopic composition of carbon ( $\delta^{13}\text{C}$ ), a character associated with water-use efficiency. To this end, an interspecific cross between *E. urophylla* x *E. grandis* was studied in four experimental trials in the Republic of Congo. Our approach, based on mapping of quantitative trait loci (QTL), combines (i) a high-throughput genotyping, (ii) a characterization of inter and intra-annual growth dynamics and  $\delta^{13}\text{C}$ , as well as a continual measurement of stem radial micro-variations and (iii) a continual characterization of environmental factors.

First of all, this work led to the construction of the first high-resolution genetic maps in *Eucalyptus*, improving the sequence of the reference genome. Then, the analysis of genetic architecture of  $\delta^{13}\text{C}$  enabled the identification of positional candidate genes, which might be involved in the variation of this trait. Finally, inter and intra-annual characterization of growth dynamics highlight that genetic architecture of adult growth is structured by responses to the environment at the juvenile stage. These responses were then analyzed using daily profiles of stem radial micro-variations, which enabled the characterization of the genetic determinants of response to the environmental factors at the juvenile stage.

Our results highlight the importance of considering growth as a dynamic trait, not only to understand its genetic basis, but also to select in a changing environment.

**Keywords:** growth dynamics, micro-dendrometer,  $\delta^{13}\text{C}$ , genetic mapping, QTL, genotype by environment interaction, *Eucalyptus*



## Avant-propos

Les recherches effectuées dans le cadre de cette thèse ont été menées au sein de l'UMR Biodiversité Gènes et Communautés (BioGeCo) à Pierroton, et de l'UMR Amélioration Génétique et Adaptation des Plantes méditerranéennes et tropicales (AGAP) à Montpellier. Ces travaux ont été menés en partenariat avec le Centre de Recherches sur la Durabilité des Plantations Industrielles (CRDPI) à Pointe-Noire, en République du Congo. Ces recherches s'intègrent dans le cadre du programme d'amélioration de l'*Eucalyptus*, mené par le Centre International de Recherche Agronomique pour le Développement (CIRAD) et le CRDPI. Ces recherches ont été financées par le département BIOS du CIRAD, pour l'allocation de recherche sur trois ans, et par deux projets européens : le projet Abiogen (FEDER) et le projet 34Joule Project (ERANET Plant KBBE). Une partie des travaux a été réalisée en collaboration avec la plateforme bio-informatique de Toulouse (GenoToul) et avec Oliver Brendel de l'Inra de Nancy.

# Remerciements

*Les remerciements, à la fois introductifs pour le lecteur et conclusifs pour l'auteur, permettent d'apercevoir les différentes dimensions d'une thèse, et sont également le reflet du travail collectif qu'elle représente.*

*Tout d'abord, je tiens à remercier le département BIOS du CIRAD pour le financement de mon allocation de recherche sur trois ans. Merci à Jean-Christophe Glaszmann, directeur de l'UMR AGAP, mon unité de rattachement, et à Rémi Petit, directeur de l'UMR BioGeCo, mon unité d'accueil.*

*Je remercie également Catherine Bastien et René Vaillancourt, d'avoir accepté d'être les rapporteurs de cette thèse, ainsi que Thierry Ameglio, Matthieu Falque, Patrice This, d'avoir accepté de faire partie du jury, car l'évaluation du travail de thèse est au moins aussi importante que le travail en lui-même.*

*Merci à Jean-Marc Bouvet mon directeur de thèse, pour l'autonomie qu'il m'a donnée dans mes recherches et pour la confiance qu'il m'a accordée.*

*L'Eucalyptus et le Congo n'étaient, avant cette thèse, que deux mots parmi d'autres. Ils me sont maintenant familiers. Pour cela, et pour m'avoir permis de faire cette thèse je remercie Jean-Marc Gion. Merci pour ton implication tout au long de cette thèse avec toujours la même conviction. Tu as su me faire partager ta vision de la recherche et tu m'as également laissé une très grande liberté dans la conduite de cette thèse. J'ai pu profiter du tabagisme passif (viens on va fumer!) durant de longues conversations sur l'analyse des données, l'écriture des articles et la thèse en général.*

*Merci également à tous mes collègues de BioGeCo, pour leur participation à ce travail ou tout simplement pour les bonnes conditions de travail à l'Artigua. Merci à la team Abiogen : Didier Bert, Régis Burlett et Raphaël Dulhoste pour leur participation à l'installation du dispositif sur Pin puis celui sur Eucalyptus. Nous sommes venus à bout de plusieurs kilomètres de câbles, de plusieurs milliers de branchements électriques ainsi que de quelques difficultés de régulation de tension. Et au-delà de l'installation, merci pour votre aide sur le suivi de l'essai ainsi que sur l'analyse de cette masse de données. Je remercie également Christophe Plomion pour ses conseils et son implication tout au long de cette thèse, sans qui le contenu aurait sans doute été moins pertinent. Merci également à Eric Mandrou pour les débuts au labo BM, pour les extractions d'ADN les plus longues jamais réalisées, puis pour le design de la puce Euca qui a donné de bons résultats, et enfin pour ta disponibilité même plus d'un an après ton départ. Merci à Marlène pour ta bonne humeur et les manips au labo avec notamment l'amélioration du protocole d'extraction de ces feuilles d'eucalyptus plutôt coriaces. Merci à Sébastien pour le rangement de tous les prélèvements de feuilles (2500) sur l'essai Abiogen, ainsi que la préparation des feuilles pour les analyses isotopiques. Merci à Patrick et à l'équipe de la plateforme (Franck, Sarah, Adeline et anciennement Thibaut et Christophe) pour m'avoir permis de réaliser mes manips dans de bonnes conditions et pour les conseils et l'aide reçus pendant ma période au labo. Merci également à Thierry, Jean-Marc et Yec'han pour leur aide sur Linux et SGE.*

*Merci enfin à mes nombreux collègues de bureau : Frank puis Emilie, Mélanie et Hélène, puis Christophe et enfin François pour la bonne ambiance de travail.*

*Je tiens à remercier l'ensemble de l'équipe GFP à Montpellier, notamment Marie Denis pour sa participation à mes comités de thèse et ses conseils en stats, ainsi que Roselyne Lannes pour sa disponibilité, son efficacité et sa gentillesse.*

*Je remercie également l'ensemble du personnel du CRDPI, à commencer par son directeur Philippe Vigneron et son directeur scientifique Aubin Saya, pour l'accueil qu'ils nous ont réservé à Pointe-Noire. Merci également à Louis Mareschal pour son aide sur place mais aussi pour les échanges de données météo et les conseils sur les sols de Pointe-Noire. Je remercie également les VI François, Gaël et plus particulièrement Blandine et Bastien pour leur accueil dans la case de passage et pour les bons moments lorsque j'étais en mission. Merci à également pour leur accueil chaleureux au CRDPI à Charles, Julienne, Crisley, Séraphin, Juste Daniel, Mélanie Prudence, Gerda, Arsène, ainsi qu'à tous les autres. J'adresse des remerciements particuliers à André Mabilia pour le travail fourni sur l'essai Abiogen et pour ton implication sans faille au cours de ces trois ans. Sans toi le dispositif Abiogen n'aurait sans doute pas aussi bien fonctionné. J'ai vraiment apprécié de travailler avec toi. Je te remercie également pour ton accueil lors de mes missions à Pointe-Noire, toujours très intenses mais sans doute parmi les meilleurs moments de ma thèse. Je me souviendrai encore longtemps du fameux sandwich au thon et à l'avocat partagé lors des repas sur le terrain.*

*Last but not least, Morgane merci pour tout. De la préparation du fameux cake (pour lequel je suis très envié) à la relecture de l'intégralité de ma thèse, en passant par la préparation des plaques d'extraction le week-end et le soutien tout au long de cette thèse. Tu as été essentielle dans la réalisation de ce document. Sans toi il n'y aurait sans doute pas eu de thèse merci de m'avoir permis de me dépasser.*

*À tous merci.*

« *Aide-toi, la science t'aidera...* »



# Sommaire

<b>Introduction.....</b>	<b>13</b>
<b>1 Contexte .....</b>	<b>17</b>
1.1 Contexte général.....	19
1.1.1 Un intérêt croissant pour les forêts plantées.....	19
1.1.2 Le changement climatique et ses incidences sur la productivité des forêts .....	21
1.1.3 Gestion et adaptation des forêts plantées .....	30
1.2 Déterminisme génétique et environnemental de la croissance chez les arbres forestiers .....	33
1.2.1 La croissance, ses relations avec les variables environnementales et la composition isotopique de carbone .....	35
1.2.2 Les bases génétiques de la croissance.....	41
1.2.3 Stratégie d'intégration de l'interaction G×E pour l'étude de l'architecture génétique de la croissance .....	45
1.3 Objectif de la thèse et démarche adoptée .....	51
Références .....	53
<b>2 Matériel et méthodes .....</b>	<b>67</b>
2.1 Matériel végétal et essais.....	69
2.2 Site d'étude et conditions pédoclimatiques .....	71
2.3 Installation, suivi et traitement du dispositif P10 .....	73
2.3.1 Conception du dispositif .....	73
2.3.2 Mise en place du dispositif.....	75
2.3.3 Amélioration et suivi du dispositif Eucalyptus .....	83
2.3.4 Traitement des données.....	85
2.4 Caractères phénotypiques mesurés .....	87
2.4.1 Croissance .....	87
2.4.2 Mesures de $\delta^{13}\text{C}$ .....	89
2.4.3 Micro-variations du rayon.....	91
2.5 Génotypage .....	91
2.5.1 Extraction d'ADN.....	93



2.5.2	Design de la puce SNP "Eucalyptus" .....	93
2.6	Cartographie génétique.....	95
2.7	Détection de QTL.....	97
	Références .....	99
<b>3</b>	<b>Plasticity of primary and secondary growth dynamics in <i>Eucalyptus</i> hybrids .....</b>	<b>103</b>
<b>4</b>	<b>High-resolution genetic linkage maps of <i>E. grandis</i> and <i>E. urophylla</i> ..</b>	<b>119</b>
4.1	Introduction.....	121
4.2	Material and Methods.....	125
4.2.1	Plant material and DNA extraction.....	125
4.2.2	SNP detection and array design.....	125
4.2.3	SNP genotyping.....	127
4.2.4	Linkage mapping .....	127
4.2.5	Linkage mapping accuracy.....	128
4.3	Results .....	128
4.3.1	Resequencing, SNP selection and SNP-array design .....	128
4.3.2	Genotyping quality .....	129
4.3.3	Genetic Mapping .....	131
4.3.4	Comparison between genetic maps and the BRASUZ1 genome .....	135
4.4	Discussion.....	138
4.4.1	Performance of the SNP Array.....	138
4.4.2	Linkage mapping accuracy.....	138
4.4.3	Validation and improvement of the <i>Eucalyptus</i> BRASUZ1 genome structure .....	140
4.4.4	Distribution and effect of SDRs .....	141
4.5	Conclusion and prospects .....	142
	References .....	164
<b>5</b>	<b>Genetic architecture of carbon isotope composition and growth in <i>Eucalyptus</i> .....</b>	<b>169</b>
5.1	Introduction.....	171



5.2	Materials and Methods .....	173
5.2.1	Mapping population and field experiments.....	173
5.2.2	Environmental characterisation .....	173
5.2.3	Phenotypic measurements .....	175
5.2.4	Linkage map construction .....	177
5.2.5	QTL analysis .....	177
5.2.6	Functional enrichment analysis in $\delta^{13}\text{C}$ QTL-regions .....	179
5.3	Results .....	179
5.3.1	Climatic conditions .....	179
5.3.2	Trait variation and correlations .....	181
5.3.3	QTL for $\delta^{13}\text{C}$ and growth traits .....	181
5.3.4	GO analysis of positional candidate genes.....	185
5.4	Discussion .....	185
5.4.1	The phenotypic variability of $\delta^{13}\text{C}$ is determined by environmental and ontogenic effects.....	185
5.4.2	Genetic architecture of $\delta^{13}\text{C}$ in eucalyptus.....	187
5.4.3	Genes underlying $\delta^{13}\text{C}$ -QTLs support differences in leaf $\delta^{13}\text{C}$ between the dry and rainy seasons .....	188
5.4.4	Relation between $\delta^{13}\text{C}$ , WUE and growth .....	190
	References .....	201
<b>6</b>	<b>Architecture génétique de la dynamique de croissance : de la croissance totale à la réponse journalière .....</b>	<b>209</b>
6.1	Introduction.....	211
6.2	Matériel et méthodes .....	213
6.2.1	Matériel végétal et essai au champ .....	213
6.2.2	Suivi des variables environnementales .....	213
6.2.3	Données phénotypiques : mesures ponctuelles et suivi en continu .....	217
6.2.4	Cartographie génétique et détection de QTL .....	219
6.3	Résultats : caractérisation de l'architecture génétique de la dynamique de croissance sur un large effectif .....	219
6.3.1	Variations et corrélations phénotypiques .....	219
6.3.2	Architecture génétique des caractères de croissance .....	223



6.4 Résultats : La plasticité des micro-variations radiales .....	225
6.4.1 Caractérisation du climat .....	225
6.4.2 Caractérisation phénotypique des paramètres du cycle circadien .....	227
6.4.3 Architecture génétique des micro-variations du rayon .....	237
6.5 Discussion .....	239
6.5.1 La croissance ponctuelle et les paramètres du cycle circadien .....	239
6.5.2 Relations entre les paramètres du cycle circadien et les variables environnementales .....	240
6.5.3 L'architecture génétique de la croissance et de la plasticité des micro-variations radiales .....	241
6.5.4 Lien entre l'architecture génétique des micro-variations et celle du $\delta^{13}\text{C}$ ainsi que celle des propriétés du bois .....	245
Figures et tableaux additionnels .....	247
Références .....	260
<b>7 Conclusion et perspectives .....</b>	<b>263</b>
7.1 Synthèse des résultats obtenus .....	264
7.1.1 Les apports des cartes génétiques haute résolution .....	264
7.1.2 Architecture génétique de la dynamique de croissance .....	265
7.1.3 Architecture génétique de la composition isotopique du carbone .....	267
7.1.4 Architecture de la plasticité des micro-variations du rayon aux variations environnementales .....	269
7.1.5 Limites associées à ces travaux .....	270
7.2 Perspectives .....	271
7.2.1 L'étude de la recombinaison .....	271
7.2.2 D'autres approches d'analyse des variations journalières du rayon .....	272
7.2.3 Le lien entre micro-variations et propriétés du bois .....	272
7.2.4 Vers la sélection/prédiction génomique .....	273
Références .....	274
<b>Annexes .....</b>	<b>279</b>





# Introduction



Une forte production de bois, résultat de la croissance apicale et radiale, est un objectif essentiel des programmes d'amélioration des arbres forestiers. Sa dynamique de formation résulte d'un large éventail de processus écophysiologiques, dont certains sont fortement dépendants des variations environnementales. C'est le cas par exemple de la photosynthèse, qui dépend notamment de la demande évaporatoire et de la disponibilité en eau du sol pour l'entrée du carbone atmosphérique dans la plante. Dans le cadre de l'évolution du climat, l'augmentation des températures moyennes combinée à un changement du régime des précipitations va accroître la fréquence et l'intensité des sécheresses, notamment en Amérique centrale, en Afrique et en Australie. Cette baisse de la disponibilité en eau devrait s'accompagner d'une réduction de la productivité des forêts plantées dans les régions concernées. Le maintien de la productivité dans un environnement plus limitant est donc un axe majeur des programmes d'amélioration des espèces forestières. Pour répondre à cet enjeu, des recherches en génétique, dont un front de science qui porte sur l'utilisation de l'information génomique pour augmenter l'efficacité de la sélection et les gains génétiques par unité de temps, sont en fort développement chez les arbres forestiers. Les travaux réalisés au cours de cette thèse sur la caractérisation de l'architecture génétique (nombre de gènes, leur localisation sur le génome et leurs effets sur la variation phénotypique) de la dynamique de croissance et de la composition isotopique du carbone de l'arbre s'intègre dans ce cadre.

Cette thèse s'inscrit également dans le cadre du programme d'amélioration génétique des eucalyptus mené conjointement par le Centre International de Recherche Agronomique pour le Développement (CIRAD) et le Centre de Recherches sur la Durabilité des Plantations Industrielles (CRDPI) en République du Congo. Depuis 1989, un schéma de sélection récurrente réciproque a été mis en place pour obtenir des combinaisons hybrides entre *Eucalyptus urophylla* et *Eucalyptus grandis*. Les premiers travaux de cartographie génétique entre ces deux espèces datent de 1992. Le croisement interspécifique utilisé lors de ces premiers travaux constitue le croisement de référence utilisé dans cette thèse. Au-delà de son importance économique, le genre *Eucalyptus* présente deux atouts pour étudier l'architecture génétique de la dynamique de croissance et son interaction avec les variables environnementales. En effet, sa croissance rapide en conditions tropicales (rotation entre 6 et 7 ans) permet un accès rapide à des arbres matures. De plus, le développement récent des ressources génomiques pour ce genre (génome de référence pour *E. grandis*, catalogues de séquences exprimées et de polymorphismes) ouvre la voie à une caractérisation fine de l'architecture



génétique de la croissance et des caractères associés, grâce à l'identification des gènes sous-jacents aux loci à effets quantitatifs (QTL).

Dans ce contexte, les recherches présentées dans ce manuscrit portent sur l'étude du déterminisme génétique et environnemental de la dynamique de croissance dans une descendance de plein-frères d'*Eucalyptus*. Elles s'appuient sur des dispositifs au champ installés en République du Congo.

Le **premier chapitre** décrit le contexte global et scientifique des travaux et aborde les concepts utilisés dans les chapitres suivants. Il présente également l'objectif de la thèse ainsi que les questions de recherche associées.

Le **second chapitre** présente le croisement biparental de référence *E. urophylla* x *E. grandis* utilisé, ainsi que les essais au champ dans lesquels le croisement a été suivi. Un focus est fait sur l'installation du dispositif de suivi en continu des micro-variations du diamètre du tronc. Les méthodes utilisées pour le traitement du signal de micro-variations, la cartographie génétique et la détection de QTL sont également décrites.

Le **troisième chapitre** traite de l'effet de la disponibilité en eau sur l'évolution temporelle du déterminisme génétique de la croissance, ainsi que de l'interaction QTL×E. Les résultats de ce chapitre proviennent de l'analyse de trois dispositifs, pour lesquels des mesures de hauteur et circonférence ont été réalisées sur 5 ans, à une échelle intra-annuelle (au cours de la saison de pluie et de la saison sèche) et interannuelle. Ce chapitre a été publié dans la revue *BMC Plant Biology* en 2013.

Le **quatrième chapitre** se focalise sur la cartographie génétique à haute résolution et son utilisation, pour mieux comprendre la structure du génome de référence de l'eucalyptus « BRASUZ1 », grâce à la comparaison entre carte physique et carte génétique. L'effet du nombre d'individus et de marqueurs sur la précision et la résolution des cartes est abordé. Ce chapitre fait l'objet d'un article préparé pour le numéro spécial "*Eucalyptus Genome*" de la revue *New Phytologist*, une fois que le génome de cette espèce sera publié. Les cartes génétiques produites servent de base pour la détection de QTL des chapitres 5 et 6.

Le **cinquième chapitre** s'attache à disséquer l'architecture génétique de la composition isotopique du carbone ( $\delta^{13}\text{C}$ ), un indicateur de l'efficacité d'utilisation de l'eau chez les plantes. L'analyse porte sur trois dispositifs, pour lesquels des mesures de  $\delta^{13}\text{C}$  ont été réalisées à partir de la cellulose extraite du bois ou des feuilles formées, en saison sèche et en



saison des pluies. Le contenu en gènes des QTL majeurs de ce caractère est étudié grâce à l'alignement des cartes génétiques sur la séquence de référence du génome. Ce chapitre est également présenté sous forme d'un article à soumettre. Il traite également de la comparaison de l'architecture génétique du  $\delta^{13}\text{C}$  et de la croissance.

Dans le **sixième chapitre**, l'analyse de l'architecture génétique de la croissance est étendue à un effectif important (960 individus). En parallèle, le signal de micro-variations du rayon sur un sous-ensemble de 222 individus est analysé à différentes échelles de temps (jour, semaine, mois). L'effet des variations environnementales sur celui-ci est décrypté. L'architecture génétique des différentes composantes de la croissance radiale est ensuite décrite.

Enfin, une conclusion générale ainsi que les perspectives de ces travaux sont abordées dans un **septième chapitre**.



# 1 Contexte

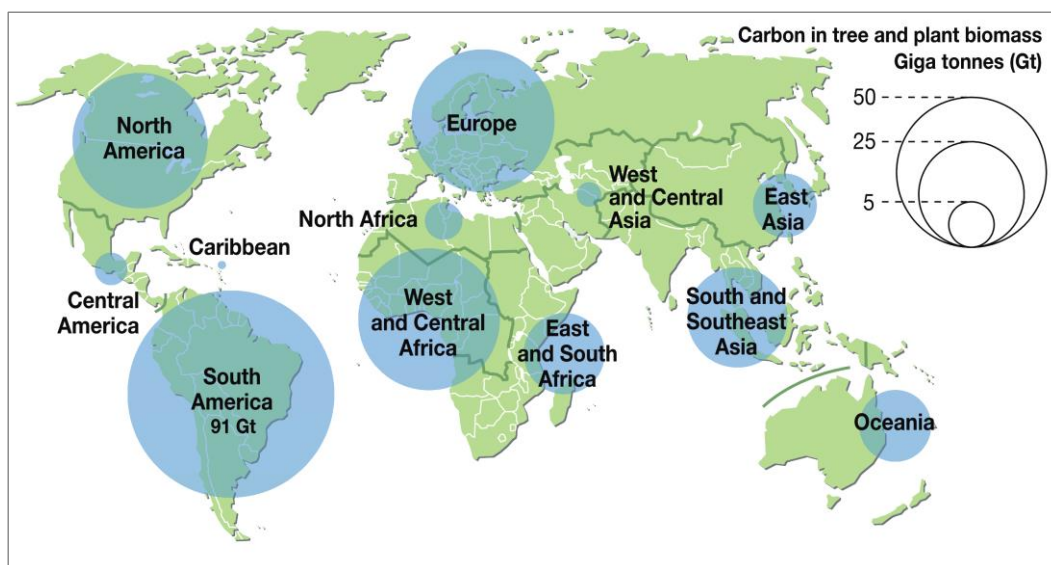


Figure 1.1 : Estimation des stocks de carbone dans les forêts par région en 2005 (d'après [1, 2]).

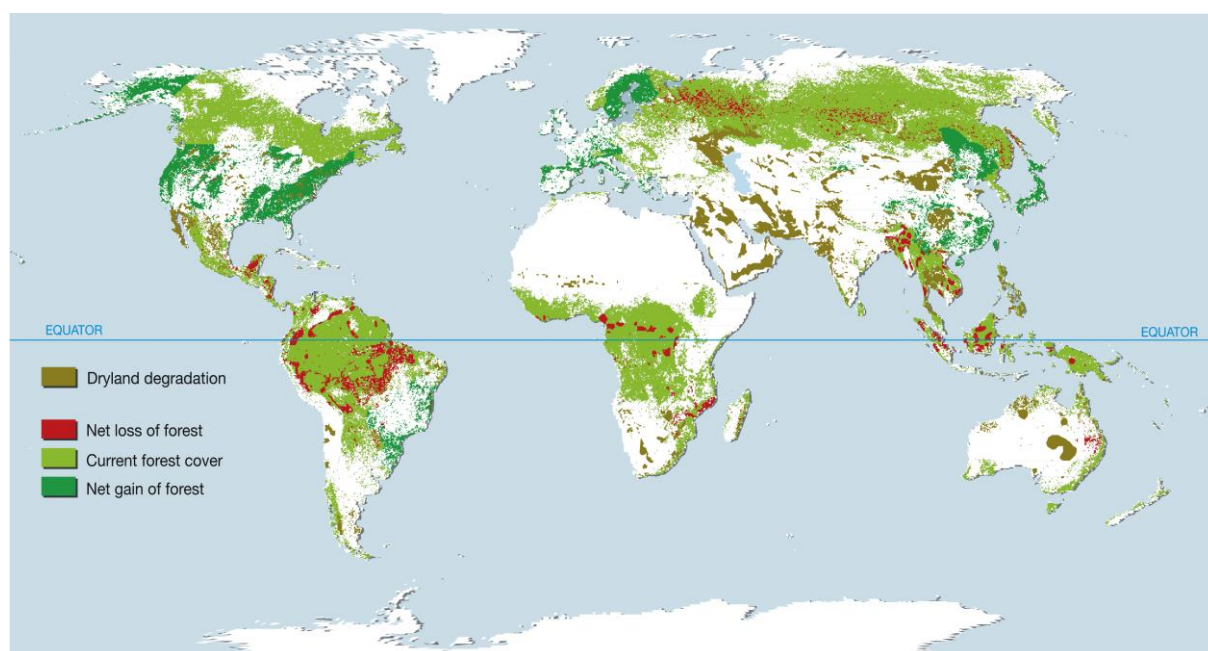


Figure 1.2 : Changements de la couverture forestière mondiale entre 1990 et 2005 (d'après [1, 3]).



## 1.1 Contexte général

### 1.1.1 Un intérêt croissant pour les forêts plantées

Les forêts ont deux rôles majeurs aux objectifs parfois contradictoires : un rôle écologique et un rôle socio-économique.

En effet, les forêts qui recouvrent actuellement environ 30% des terres émergées, soit quatre milliards d'hectares [4], sont considérées comme le réservoir de 80% de la biodiversité terrestre. Ainsi, 12% des forêts du monde sont principalement assignées à la conservation de la diversité biologique [5, 6]. Par ailleurs, les forêts contribuent également à la protection des sols contre l'érosion [7], à la lutte contre la désertification et à la préservation des ressources en eau. Enfin, les forêts jouent un rôle essentiel dans le cycle du carbone, d'une part en fixant une partie du carbone atmosphérique par photosynthèse et d'autre part parce qu'elles représentent environ 45% du carbone terrestre (652 Gt) contenu majoritairement dans la biomasse (289 Gt) et dans les sols (291,5 Gt, Figure 1.1) [5]. Compte tenu de leurs différentes fonctions, les forêts sont donc au centre des préoccupations environnementales sur l'évolution du climat (réduction de l'émission de gaz à effet de serre) [8] et sur la conservation de la biodiversité (réduction de la déforestation).

Parallèlement, 30% des forêts de la planète sont affectées en premier lieu à la production de bois et 24% sont destinées à des usages multiples, comprenant la production de bois et de produits forestiers non ligneux (fruits, champignons, gibiers) [5]. Environ la moitié du bois récolté correspond à du bois de feu, l'autre moitié à du bois rond à usage industriel (bois de trituration, bois d'œuvre,...). La production de bois au niveau mondial représentait en 2005, 3,4 milliards de m<sup>3</sup> soit 0,7% de la totalité du matériel sur pied. La valeur des extractions de bois était estimée à 97 milliards de dollars US. Quant à la valeur totale des produits non ligneux extraits des forêts en 2005, elle se chiffrait à 18,5 milliards de dollars US. Outre la valeur marchande de produits forestiers, les différentes étapes de la chaîne de production du bois emploient 10,5 millions de personnes dans le monde, majoritairement en Asie [4]. On estime qu'environ 1,2 milliard de personnes dépendent de la forêt pour subsister [4, 9].

L'exploitation importante du bois et la conversion des forêts en surfaces agricoles, conjuguées à des phénomènes naturels (feu, ravageur,...), tendent à diminuer les surfaces forestières, notamment celles des forêts naturelles [5] (Figure 1.2). Bien que cette tendance ne soit pas récente [10], elle s'est accélérée au cours des dernières décennies avec une

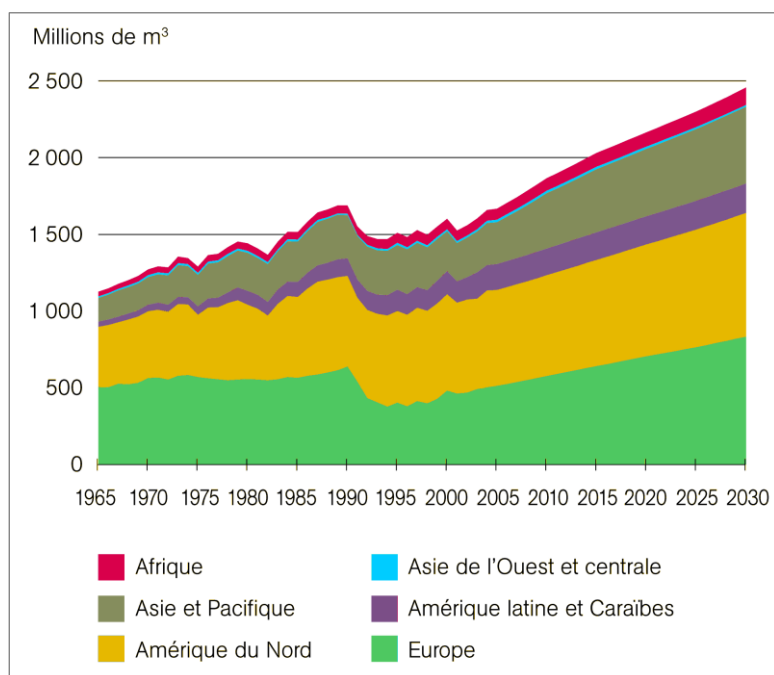


Figure 1.3 : Production mondiale de bois rond industriel (issu de [11])

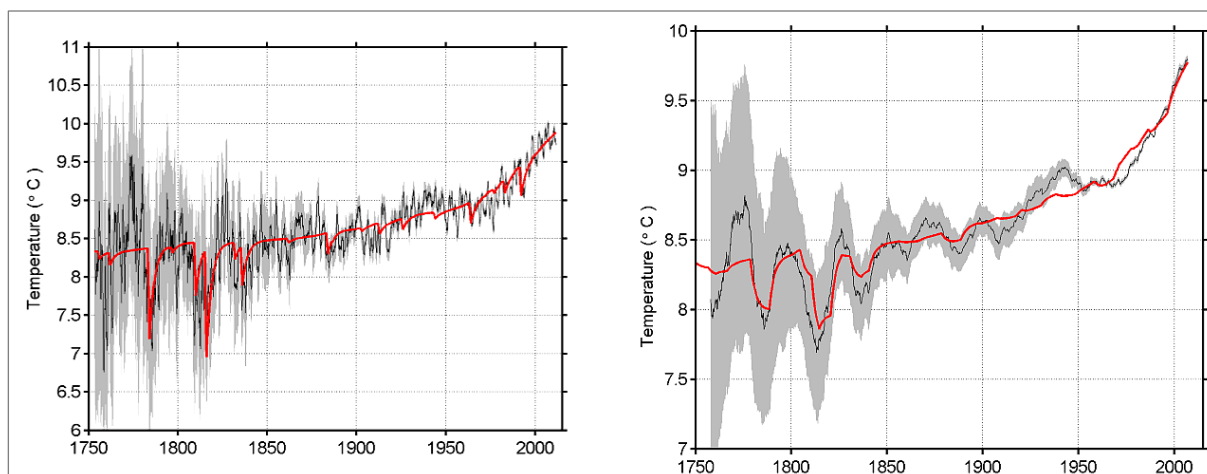


Figure 1.4 : Evolution des températures moyennes à la surface globe: à l'échelle annuelle (à gauche) et décennale (à droite). Les données issues de l'étude de Berkeley Earth sont en noir et celles issues d'un modèle utilisant les émissions volcaniques de sulfate et la teneur atmosphérique en  $\text{CO}_2$  sont en rouge. La teneur en  $\text{CO}_2$  est responsable de la montée progressive et le sulfate d'origine volcanique est responsable des courtes réductions. L'inclusion de l'activité solaire n'améliore pas de manière significative l'ajustement du modèle. La zone grise est l'intervalle de confiance à 95%. (d'après [12])





diminution de 13 millions ha/an de 2000 à 2010. Ces pertes ne sont qu'en partie compensées par l'augmentation des surfaces de forêts plantées (+ 4,9 millions ha/an de 2000 à 2010) [4] et l'expansion naturelle des forêts (+ 2.9 millions ha/an de 2000 à 2010), donnant des pertes nettes évaluées à environ 5.2 millions ha/an entre 2000 et 2010 [5]. L'augmentation de la demande en bois, poussée par la croissance démographique mondiale et les politiques environnementales, incitant à augmenter la part des énergies et des matériaux renouvelables tels que le bois, devrait accentuer cette tendance au cours des prochaines années. Ainsi, la consommation de bois rond à usage industriel est estimée à 2,44 milliards de m<sup>3</sup> pour 2030 (Figure 1.3), contre 1,68 milliards de m<sup>3</sup> en 2005 [5, 11]. En 2005 les forêts plantées satisfaisaient 66% de la production de bois rond industriel [13]. Les forêts plantées représentaient 264 millions d'hectares en 2010, soit 6,6 % des surfaces de forêts, et devraient atteindre 300 millions ha d'ici 2020. Cette hausse de la superficie devrait permettre de satisfaire la majeure partie de la demande en bois, atténuant ainsi la pression qui s'exerce sur les forêts naturelles [14].

Les forêts plantées sont composées aux trois quarts par des essences indigènes, le reste correspondant à des espèces introduites [5]. La majorité des surfaces plantées le sont en conifères (54% en 2005), le genre *Pinus* représentant à lui seul 32% des surfaces. Pour les feuillus (39% en 2005), le genre *Eucalyptus* est majoritaire avec 8% des surfaces. Il est essentiellement planté en régions tropicales, avec d'importantes surfaces en Inde, au Brésil et en Chine qui représentent 56% des plantations d'*Eucalyptus*. Certaines espèces plus résistantes au froid sont plantées en régions tempérées. En république du Congo, les plantations industrielles de la région de Pointe-Noire représentent la quasi-totalité des boisements d'*Eucalyptus* du pays soit environ 36 000 ha.

### 1.1.2 Le changement climatique et ses incidences sur la productivité des forêts

Le climat de la Terre est un système complexe en constante évolution, sous l'influence de différents facteurs internes (composition atmosphérique, volcanisme,...) et externes (rayonnement solaire, rotation de la Terre,...) qui agissent à différentes échelles de temps et d'espace sur les différents compartiments de la terre (atmosphère, hydrosphère, biosphère). Les changements climatiques récents, observés au cours des dernières décennies montrent une accélération de ces évolutions, avec notamment une augmentation de la température moyenne de la surface terrestre de 0.6 à 0.9 °C sur les 50 dernières années [12, 15-19] (Figure 1.4).

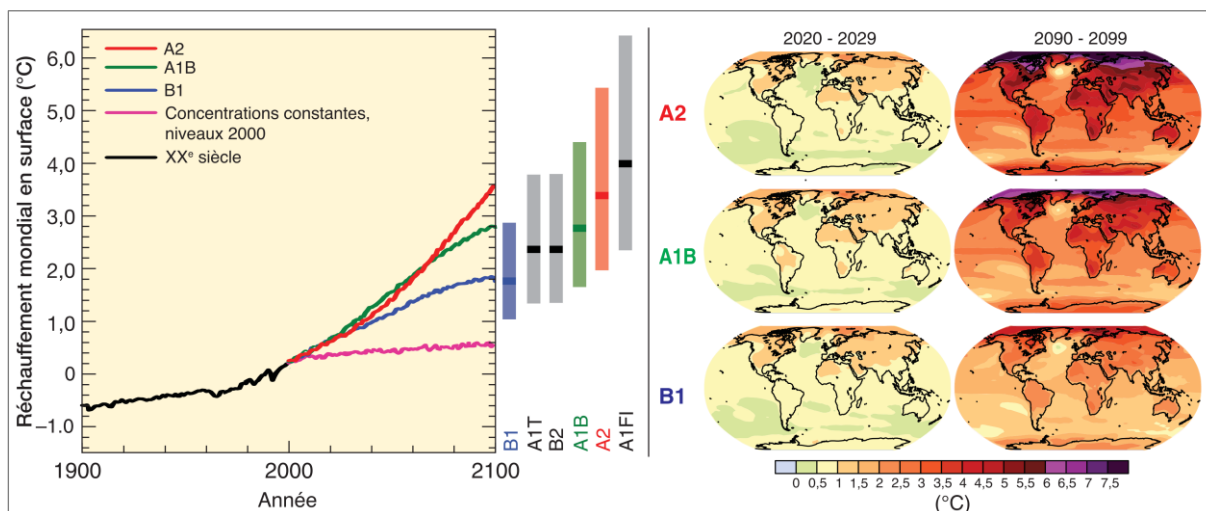


Figure 1.5 : **À gauche** : Évolution des températures moyennes à la surface du globe par rapport à la période 1980-1999. Différents scénarios sont représentés (A2, A1B et B1). Ces scénarios sont basés sur la production de GES selon différentes hypothèses d'évolutions des sociétés [16]. Les barres au milieu de la figure indiquent les valeurs les plus probables (zone foncée) et les fourchettes probables selon les six scénarios étudiés. **À droite** : Évolution projetée de la température en surface pour le début et la fin du XXI<sup>e</sup> siècle par rapport à la période 1980-1999, pour les scénarios A2 (en haut), A1B (au milieu) et B1 (en bas). (d'après [16])

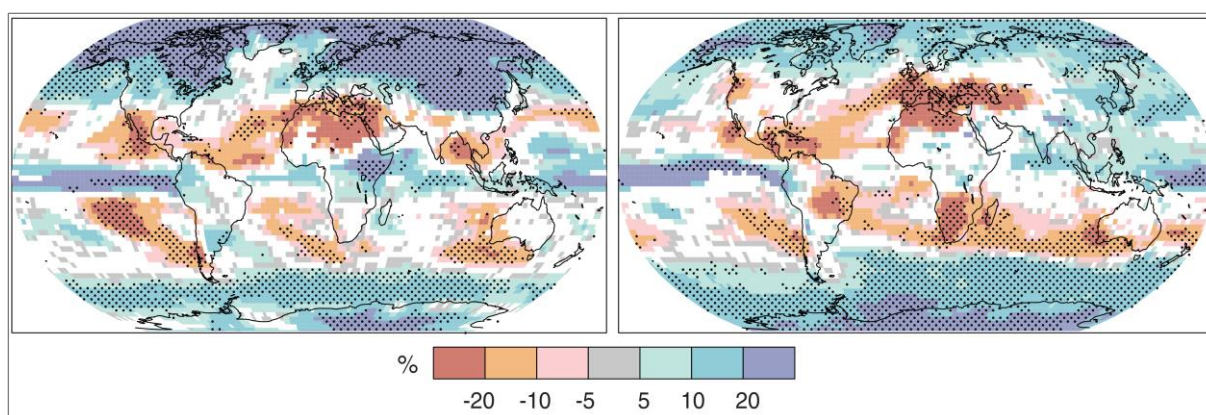


Figure 1.6 : Variations relatives du régime des précipitations (%) pour la période 2090-2099, par rapport à la période 1980-1999. Les valeurs indiquées sont des moyennes tirées de plusieurs modèles pour des périodes allant de décembre à février (à gauche) et de juin à août (à droite). Les zones en blanc correspondent aux régions où moins de 66 % des modèles concordent sur le sens de la variation, et les zones en pointillés à celles où plus de 90 % des modèles concordent sur celui-ci. (d'après [16])



L'impact des activités humaines (production de gaz à effet de serre (GES), notamment le CO<sub>2</sub>) sur ces changements a clairement été mis en évidence par différentes études [12, 15-17] (Figure 1.4). Compte tenu de la progression constante des émissions de GES au niveau mondial, avec une prévision de 59 Gt équivalent CO<sub>2</sub>/an en 2020 [20], et de la longue durée de vie du CO<sub>2</sub> dans l'atmosphère, l'augmentation des températures moyennes devrait donc se poursuivre. Les différents modèles prévoient une hausse globale entre 1 et 2°C pour la période 2020-2029, et entre 1 et 3,7°C pour la période 2090-2099, par rapport à la période 1980-1999 (Figure 1.5). Cette hausse des températures va entraîner une augmentation de l'évaporation et des précipitations de façon hétérogène à la surface du globe [21]. Les latitudes élevées recevront des précipitations plus importantes en raison de la capacité supplémentaire de transport d'eau de la troposphère. De nombreuses régions arides et semi-arides des latitudes moyennes verront une diminution des précipitations. Les changements les plus importants de précipitations portent sur le nord de l'Eurasie et de l'Amérique du Nord, et devraient survenir surtout en hiver (Figure 1.6). Contrairement à l'augmentation de la concentration atmosphérique en CO<sub>2</sub> qui sera globalement homogène à la surface du globe, les températures et les précipitations ne changeront pas de manière uniforme, impliquant une variabilité régionale forte.

Malgré les incertitudes liées aux scénarios et aux modèles de projection, sur l'intensité, le rythme et la répartition spatiale des changements climatiques [16], il est désormais certain que l'évolution du climat affectera toute la biosphère. Les forêts, qu'elles soient naturelles ou plantées, seront touchées par l'évolution du climat [22], notamment par l'augmentation de la concentration atmosphérique en CO<sub>2</sub>, l'augmentation des températures et le changement de la distribution des précipitations. Par ailleurs, sous l'effet du changement climatique, les dynamiques des bio-agresseurs risquent d'être modifiées et les phénomènes extrêmes, tels que les feux de forêts ou les tempêtes risquent d'être plus fréquents. L'effet de ces différents facteurs sur la productivité des forêts est détaillé par la suite. Il est important de noter qu'ils interagissent bien qu'ils soient présentés séparément.

#### *1.1.2.1 Effet de l'augmentation de la concentration en CO<sub>2</sub>*

La concentration atmosphérique actuelle en CO<sub>2</sub> (≈ 385 ppm) ne permet pas à la RuBisCo (enzyme qui permet la fixation du CO<sub>2</sub>) de fonctionner en condition saturante en CO<sub>2</sub>. En effet, elle sature lorsque la concentration en CO<sub>2</sub> tend vers 1000 ppm, ce qui est supérieur





aux prévisions de concentration atmosphérique en CO<sub>2</sub> pour 2100 ( $\approx$  700 ppm). L'augmentation de la concentration en CO<sub>2</sub> permettra donc, pour les plantes en C<sub>3</sub> (tous les arbres), une meilleure activité photosynthétique, une réduction de la conductance stomatique et donc une meilleure efficacité d'utilisation de l'eau (WUE) [23-25]. Des expériences d'enrichissement en CO<sub>2</sub> (FACE = free-air CO<sub>2</sub> enrichment, OTC = open-top chamber) à moyen et long termes ont montré une augmentation de la croissance des arbres [24-27]. Cette tendance générale masque des différences parfois fortes. En effet, en relation avec plusieurs facteurs : la région, l'espèce, l'âge et la durée de l'expérience, les gains en croissance peuvent être nuls [28-31] ou s'élever jusqu'à 23% (en moyenne) sur *Pinus taeda*, *Liquidambar styraciflua* et *Populus tremuloides*, *alba*, *nigra* [32], 27% sur *Pinus taeda* [33], ou 33% sur *Larix decidua* [29]. Outre les effets sur la croissance, l'élévation de la concentration en CO<sub>2</sub> peut avoir un effet sur les propriétés du bois comme la densité, la taille des vaisseaux ou le pourcentage de cellulose [34-36].

Ces études FACE sur les arbres forestiers ont été conduites en régions tempérées ou boréales, mais à ce jour aucune expérience n'a été conduite en régions tropicales, bien que la compréhension des effets d'une augmentation de CO<sub>2</sub> sur les forêts tropicales soit cruciale [37, 38]. Toutefois, des modèles ont été développés afin de prédire l'effet d'une augmentation de CO<sub>2</sub> dans ces régions [39]. Par exemple, Almeida *et al.* [40] ont montré une augmentation de 26% de la productivité sur plantation d'*Eucalyptus* au Brésil, à l'horizon 2050. Une étude FACE sur *Eucalyptus* a débuté en 2012 dans la région de Sydney en Australie [41], et devrait permettre de mieux comprendre et prédire les effets d'une augmentation de la concentration en CO<sub>2</sub> sur ce type de végétation.

Les gains de productivité observés sont néanmoins plus faibles que les gains sur l'activité photosynthétique [42, 43]. Différentes hypothèses ont été formulées pour expliquer cette différence. Une acclimatation peut survenir après une exposition sur le long terme à une concentration élevée en CO<sub>2</sub>, avec notamment une réduction de l'activité et de la quantité des enzymes photosynthétiques (RuBisCo) [44]. Le transfert vers les puits de carbone de l'arbre (racine, tronc) étant limité, cette acclimatation peut être reliée à une accumulation de carbohydrates dans les feuilles, ce qui entraîne une régulation négative de la photosynthèse [45]. Néanmoins, différentes espèces comme le peuplier [46] ou le chêne sessile [47] ne présentent pas d'acclimatation sur le long terme. Une autre hypothèse est la limitation de la

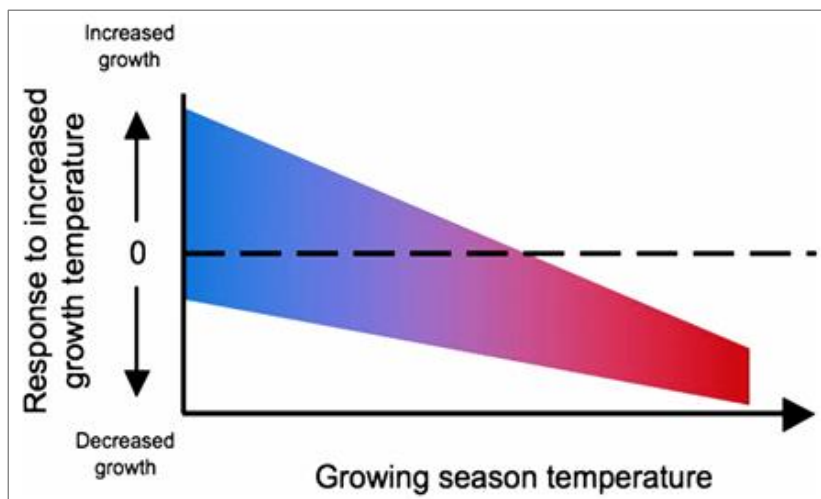


Figure 1.7 : Représentation schématique de la réponse des arbres à une augmentation des températures en fonction de leur localisation géographique. La région colorée indique la fourchette de réponse. Les espèces provenant d'un environnement froid (en bleu) sont susceptibles d'avoir une meilleure croissance si l'eau et les nutriments ne sont pas limitant. Les espèces provenant d'un environnement chaud (en rouge) devraient présenter une diminution de la croissance avec une variabilité moins importante entre géotypes ou entre espèces (D'après [48])



productivité par d'autres facteurs jouant un rôle dans la croissance, comme la disponibilité en eau et en nutriments [25, 27, 42].

### 1.1.2.2 Effet de l'augmentation des températures

La température ayant un rôle dans la plupart des processus biologiques et chimiques, son augmentation aura notamment un effet sur la photosynthèse, la transpiration et la respiration [48-50], ce qui pourrait avoir un effet positif sur la croissance. De plus, l'augmentation des températures moyennes aura des conséquences sur la phénologie foliaire et/ou cambiale [51-54], conduisant un allongement de la saison de croissance. Enfin, la minéralisation du carbone et de l'azote du sol et de la litière devrait être accélérée par la hausse des températures [55, 56].

Une hausse des températures se traduira différemment en termes de croissance selon la latitude, les espèces considérées et l'interaction avec d'autres facteurs, notamment la disponibilité en eau et en nutriments [56, 57]. Pour les hautes latitudes ou altitudes, la température est un facteur limitant pour la croissance. Une hausse des températures devrait permettre, au moins à court terme, une augmentation de la productivité en améliorant la photosynthèse et en allongeant la période de croissance [56, 58]. Par exemple, Rossi *et al.* [53] ont montré qu'une augmentation des températures moyennes allongera la période de croissance du xylème de 8 à 11 jours/°C sur *Picea mariana*. Néanmoins, une plus grande variabilité de réponse à une augmentation des températures est attendue pour les espèces de ces régions, comparées à des espèces tropicales ou de climats chauds (Figure 1.7). En effet, l'augmentation des températures peut avoir un effet neutre ou négatif selon les espèces étudiées [48, 59, 60]. À l'opposé, les espèces de climats plus chauds seront en majorité pénalisées par la hausse des températures. Différentes études ont montré que la croissance de différentes espèces d'arbres est négativement corrélée avec les températures minimales dans les régions chaudes [61, 62].

Ces résultats, de l'effet à long terme d'une hausse des températures sur la croissance, sont issus de l'analyse des relations entre le climat passé et la croissance, mais aussi d'approches expérimentales d'augmentation de la température au champ, ces dernières ayant tendance à sous-estimer les effets de la température. Ces deux types d'approches tendent à donner des résultats convergents [63].



**Encadré 1 : L'eau dans la plante**

L'eau absorbée par la plante sert à assurer des fonctions physiologiques majeures comme [64]:

- Le transport de solutés (sève brute et élaborée) au niveau cellulaire et entre organes, et le refroidissement par transpiration,
- La photosynthèse, où l'eau intervient comme donneur d'électrons, et d'autres réactions métaboliques où l'eau sert de réactif ou de substrat.
- La turgescence et la croissance des différents organes.

La majeure partie (99%) de l'eau absorbée par la plante ne fait que transiter par celle-ci. L'eau passe du sol au végétal (via les poils absorbants) et du végétal vers l'atmosphère (via les stomates) constituant ainsi le continuum sol-plante-atmosphère [65]. Cette circulation s'effectue en suivant un gradient de potentiels hydriques décroissants (théorie de la tension-cohésion [66]). La transpiration, en abaissant le potentiel hydrique des parties aériennes crée une différence de potentiel entre le système aérien et le système racinaires ce qui permet à l'eau d'être absorbée au niveau des racines puis de circuler dans la plante. Afin de maintenir son statut hydrique, la plante régule ce flux grâce aux stomates, qui en fonction de leur degré d'ouverture, opposent une résistance plus ou moins forte aux échanges gazeux entre la plante et l'atmosphère. Au niveau de ses stomates la plante perd de l'eau, mais fait également rentrer du CO<sub>2</sub> nécessaire à la photosynthèse. La transpiration et la photosynthèse sont donc deux mécanismes intrinsèquement liés.

En conditions de déficit hydrique (quantité d'eau transpirée supérieure à la quantité d'eau absorbée) le métabolisme de la plante est perturbé, ce qui provoque notamment une réduction de la croissance et du développement des organes [67-69]. Un stress hydrique peut être dû à deux facteurs distincts :

- Une forte demande évaporatoire à un instant donné (échelle de temps courte).
- Une disponibilité en eau du sol insuffisante (échelle de temps plus longue).

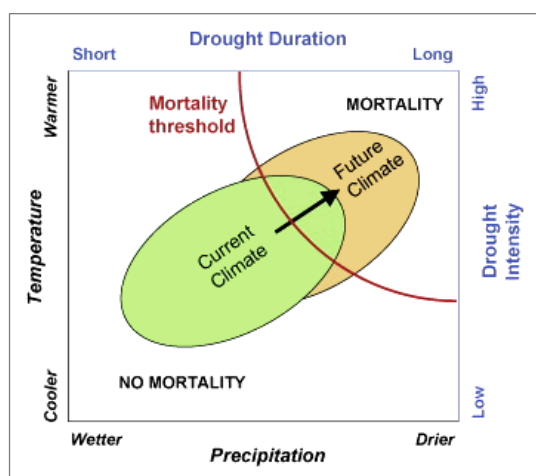


Figure 1.8 : Représentation schématique de l'effet de l'évolution du climat sur la mortalité des arbres. Le seuil de résistance à la sécheresse ne variant pas, la diminution des précipitations et l'augmentation des températures vont conduire à des dépassements plus fréquents de ce seuil, et donc à une mortalité plus élevée. D'après [70]





### 1.1.2.3 Effet du changement des précipitations

Parmi les changements climatiques prévus, l'évolution des précipitations (volume et répartition) est celui qui aura la plus grande incidence sur la productivité des forêts naturelles et plantées. En effet, l'eau est un constituant cellulaire majeur et joue un rôle dans de nombreux processus physiologiques (Encadré 1). Le volume des précipitations, ainsi que leur distribution au cours de l'année, conditionnent la disponibilité en eau pour la plante et ont une influence forte sur la croissance des arbres [71-77]. Pour des plantations d'*Eucalyptus* au Brésil, suivant un gradient de pluviométrie, la production de bois varie entre 12 t/ha/an en moyenne pour les sites les plus secs (897 mm) et 24 t/ha/an pour les sites les plus humides (1411 mm) [73]. Par ailleurs, il a été montré qu'une part importante des forêts du globe sont sensibles aux sécheresses, entraînant une mortalité pouvant aller jusqu'à 100% dans certains cas [70, 78]. Cette sensibilité des arbres forestiers à la sécheresse peut s'expliquer par un fonctionnement proche de leur limite hydraulique en conditions contraintes [79, 80]. En effet, un stress hydrique peut provoquer une embolie gazeuse (cavitation) dans le xylème, ce qui réduit la capacité des arbres à fournir de l'eau aux feuilles pour les échanges gazeux photosynthétiques et peut entraîner la mort. La réduction des précipitations dans plusieurs régions [21] du monde va accentuer les risques de sécheresse, et donc entraîner une diminution de la croissance des arbres et une augmentation de la mortalité (Figure 1.8).

Les propriétés du bois peuvent elles aussi être modifiées par la disponibilité en eau. Wimmer et al. [81] ont montré une réduction de l'angle des microfibrilles lorsque les arbres subissent un stress hydrique. La densité du bois est également modifiée selon la disponibilité en eau : elle augmente avec le déficit hydrique [82, 83]. D'autres propriétés, comme le pourcentage de lignine ou de cellulose peuvent varier. Suivant l'utilisation finale du bois (bois d'œuvre, pâtes à papier, charbon), ces variations des propriétés du bois peuvent-être importantes pour les procédés de transformation.

### 1.1.2.4 Effet des événements extrêmes et des bio-agresseurs

Bien que leur prévision ne soit pas aisée du fait de leur caractère aléatoire, les phénomènes extrêmes, comme les tempêtes ou les feux, augmenteraient avec l'évolution du climat [84-86]. Flannigan *et al.* [87] ont modélisé, en relation avec le changement climatique, une augmentation de 74 à 118% de la superficie brûlée au Canada à la fin du 21<sup>ème</sup> siècle. Ces événements extrêmes peuvent avoir une incidence importante sur la productivité des forêts et



des conséquences néfastes au-delà du seul domaine forestier [88]. Les feux de forêts représentaient environ 350 millions d'hectares de pertes dans le monde en 2000. Le vent cause lui aussi des dommages importants : les tempêtes de décembre 1999 en France avec 140 millions de m<sup>3</sup> de bois endommagés en sont un exemple.

Par ailleurs, il est généralement admis que le changement climatique est susceptible d'augmenter l'activité de différentes espèces de ravageurs et de pathogènes, et de modifier leurs aires de répartition [89-91]. Les températures plus chaudes ont déjà amélioré les possibilités de propagation des insectes thermophiles [92]. Les conséquences varient d'une défoliation et perte de croissance ponctuelles à des dommages à l'échelle de l'ensemble du massif forestier. Toutefois, la modélisation des impacts du changement climatique sur les épidémies d'insectes et de pathogènes reste limitée [93, 94].

### 1.1.3 Gestion et adaptation des forêts plantées

Les forêts plantées sont par définition le produit de l'intervention humaine. Contrairement aux forêts naturelles, une adaptation planifiée face au changement climatique est donc plus facilement envisageable [95]. La connaissance des effets de l'évolution du climat sur la productivité des forêts plantées est un prérequis indispensable pour réduire leur vulnérabilité face au changement climatique. Différentes échelles d'action (massif forestier, bassin versant, parcelle, arbre) peuvent être combinées, afin d'apporter des solutions aux différents risques, mais également pour tirer parti de certains bénéfices du changement climatique.

La gestion à grande échelle (massif) est plutôt associée à la gestion de la productivité des forêts plantées sur le long terme [96]. Des moyens sont déjà mis en œuvre pour adapter le matériel forestier de reboisement en fonction des potentialités des sites. Les principaux critères pris en compte sont la production et la qualité du bois, la résistance aux bio-agresseurs et la tolérance au déficit hydrique et au froid (notamment dans le cas de l'eucalyptus) [97, 98]. L'évolution du climat nécessite de prendre en compte les contraintes futures, grâce à des modèles de production intégrant l'effet de la réduction de la disponibilité en eau [40] ou des dynamiques de ravageurs [93]. Cette anticipation des risques permet de privilégier une espèce ou un génotype plutôt qu'un autre dans des conditions changeantes. Le changement climatique n'a pas que des effets néfastes. Par exemple l'augmentation des températures moyennes permet d'envisager l'utilisation d'espèces ou de génotypes plus sensibles au froid mais plus productifs.



La gestion de la ressource en eau passe également par une gestion à l'échelle du bassin versant, afin de minimiser l'impact des changements de couverts (coupe rase, boisement) sur le cycle hydrologique [99, 100]. L'adaptation à grande échelle doit également tenir compte des prévisions d'augmentation des risques d'incendies et d'attaques de bio-agresseurs, afin de structurer le massif pour minimiser les pertes [101]. La mise en place de ces pratiques de gestion planifiée à grande échelle peut cependant être limitée par l'acquisition des données et le suivi régulier nécessaire pour mener à bien ce type de gestion.

À l'échelle de la parcelle, c'est l'ensemble de l'itinéraire technique qui peut être adapté. Les effets des pratiques sylvicoles (préparation du sol, densité de plantation, fertilisation initiale, élagage, éclaircies, irrigation,...) sur la croissance à moyen terme sont bien connus et ont permis des gains de productivité importants au cours des dernières décennies [97, 102-106]. Par contre, les interactions entre ces pratiques et les contraintes environnementales sont encore peu connues, sauf pour la contrainte hydrique [107-109]. En effet, différentes études se sont focalisées sur les pratiques sylvicoles permettant de réduire l'impact d'une sécheresse sur la croissance, notamment par une diminution de la densité de plantation ou par des éclaircies [107].

Quelle que soit l'échelle de gestion, la sélection variétale est un élément central pour l'adaptation des forêts plantées au changement climatique. En effet, le maintien de la productivité par l'adaptation des pratiques sylvicoles présente des limites. La durée de rotation importante des espèces forestières favorise leur exposition aux contraintes environnementales changeantes. Les rotations pour des espèces à croissance rapide restent relativement longues : 25 à 30 ans pour le *Pinus radiata*, 7 à 10 ans pour les hybrides *E. grandis* x *E. urophylla*. Les variétés améliorées présentent généralement de bonnes performances sur plusieurs caractères d'intérêts [97], c'est pourquoi elles présentent un avantage important dans les systèmes de production que sont les forêts plantées. Dans ce contexte, l'amélioration génétique pour la croissance en condition contrainte est incontournable pour le maintien de la productivité à moyen et long terme.

**Encadré 2 : Principe fondamental de la génétique quantitative.**

La génétique quantitative se base sur la décomposition du phénotype (P) d'une population donnée en une somme des effets génétiques (G) et de l'effet de l'environnement (E) :

$$P = G + E$$

Si on s'intéresse à la variance phénotypique ( $V_P$ ) au sein d'une population, elle aussi peut-être décomposé par la somme de la variance génétique ( $V_G$ ) et de la variance environnementale ( $V_E$ ) ainsi que 2 fois covariance entre G et E. En expérimentation, la répartition aléatoire des génotypes dans l'environnement permet de considérer l'indépendance entre G et E et donc leur covariance comme nulle. On obtient ainsi [110].

$$V_P = V_G + V_E$$

On peut ensuite décomposer  $V_G$  (sur la base du modèle infinitésimal [111]) en variance additive ( $V_A$ ), variance de dominance ( $V_D$ ), et variance épistatique ( $V_I$ ) :

$$V_G = V_A + V_D + V_I$$

L'héritabilité au sens large ( $H^2$ ) donne la part de variance phénotypique d'origine génétique :

$$H^2 = \frac{V_G}{V_P}$$

L'héritabilité au sens strict ( $h^2$ ) caractérise la part de variance phénotypique qui se transmet (variance additive) à la descendance :

$$h^2 = \frac{V_A}{V_P}$$

L'héritabilité ( $H^2$ ,  $h^2$ ) varie ainsi théoriquement entre 0 et 1. Plus l'héritabilité d'un caractère est proche de 1, plus le déterminisme génétique de ce caractère est important. Inversement un caractère qui ne dépend pas du génotype des individus a une héritabilité proche de 0.



## 1.2 Déterminisme génétique et environnemental de la croissance chez les arbres forestiers

La sélection vise à optimiser les valeurs d'un ou plusieurs caractères d'intérêt, grâce à la variabilité génétique naturelle présente au sein des espèces. Chez les arbres forestiers, cette variabilité génétique est importante du fait de leur régime de reproduction préférentiellement allogame et de leur faible domestication [112, 113]. Les caractères cibles de la sélection varient selon les espèces et le contexte environnemental et socio-économique, mais les principaux sont : la croissance, les propriétés du bois (densité, % de lignine, % de cellulose) et la résistance aux bio-agresseurs (ex : rouilles, chancres, ... [114]). Les stratégies de sélection dépendent de la biologie de l'espèce, de la variabilité et de l'héritabilité du caractère étudié (Encadré 2) et éventuellement de la possibilité d'hybridation entre espèces phylogénétiquement proches. Cependant, quelle que soit l'espèce, il existe des contraintes intrinsèques à la sélection chez les arbres forestiers : i/ L'espace nécessaire à un programme d'amélioration est généralement important. Les différents tests (provenance, descendance, clonal,...) au champ couvrent souvent plusieurs dizaines d'hectares. De plus, la mise en place d'une pépinière est indispensable pour l'élevage des plants. ii/ Le temps nécessaire pour obtenir les phénotypes des caractères d'intérêt économique est relativement long, car ce sont majoritairement ceux de l'arbre adulte (croissance, propriétés du bois). En effet les corrélations juvénile-adulte pour ces caractères sont relativement faibles [115-117], et ne permettent pas une efficacité de la sélection satisfaisante avant le tiers de la durée de rotation en général. Pour les espèces d'eucalyptus plantées dans les tropiques, la sélection s'opère à partir de 3-4 ans. Elle est de l'ordre de 10-12 ans pour le pin maritime en France. Par ailleurs, même si des méthodes favorisant la floraison précoce ont été mises au point (notamment chez l'eucalyptus, [118, 119]) plusieurs années sont souvent nécessaires avant la maturité sexuelle et donc l'obtention de descendants, que ce soit par croisement contrôlé ou non. iii/ L'effet de l'environnement est relativement important sur les caractères d'intérêt économique du fait de la pérennité des espèces. L'héritabilité au sens strict pour ces caractères est en général faible à moyenne (0,1 à 0,6) [120, 121] et varie en fonction de l'âge [122, 123].

Afin de réduire ces contraintes, des recherches utilisant des prédictors précoces moléculaires sont en cours, pour améliorer la connaissance du fonctionnement du génome des

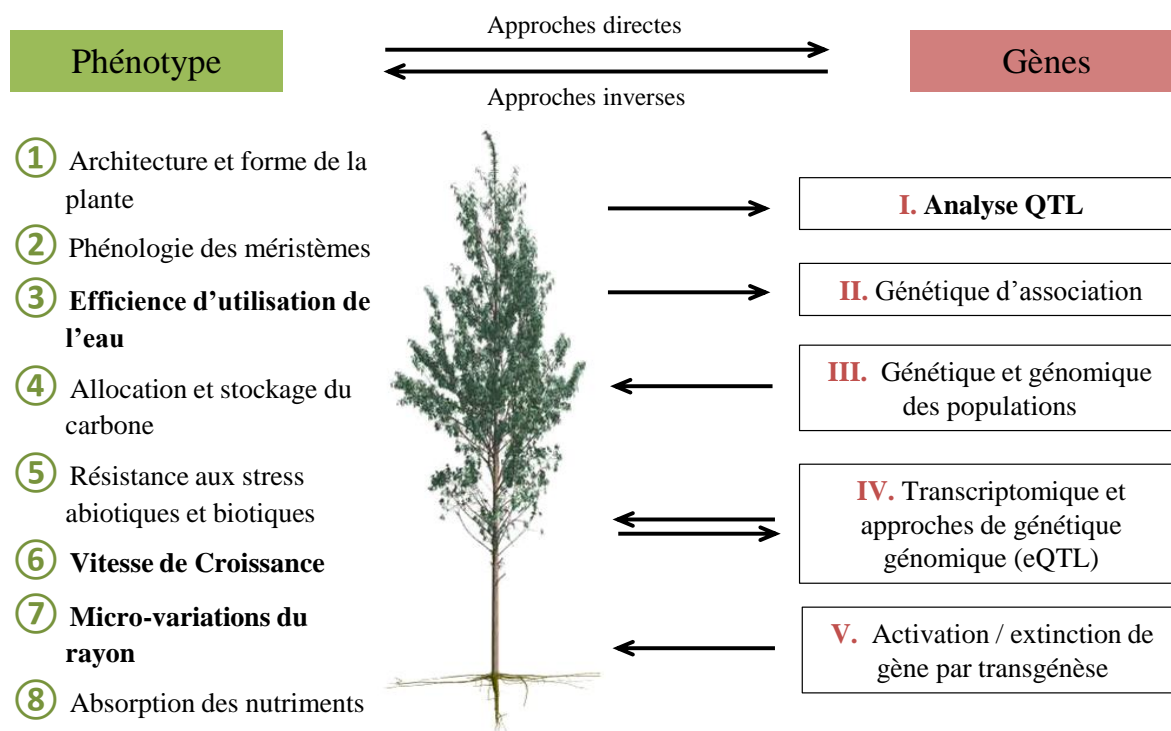


Figure 1.9 : Principaux caractères en relation avec la croissance des arbres et approches génomiques utilisées pour les étudier. Les flèches indiquent le sens de l'approche : génétique directe (du phénotype au gène) vs génétique inverse (du gène au phénotype). Sont indiqués en gras les caractères étudiés et l'approche utilisée dans le cadre de cette thèse. Adapté de [124].



arbres forestiers [125-129]. L'objectif de ces recherches est, d'améliorer l'efficacité de la sélection en augmentant les gains génétiques par unité de temps grâce à la sélection assistée par marqueurs au sens large. Cela passe par l'identification des bases génétiques et environnementales des caractères d'intérêt [124-126]. Les différentes méthodes utilisées pour y parvenir sont présentées dans la Figure 1.9. Dans cette thèse, la méthode utilisée pour identifier les bases génétiques de la croissance est l'analyse de liaison dans une descendance de plein-frères entre la variabilité phénotypique et le polymorphisme génétique au niveau de marqueurs moléculaires (analyse QTL).

### 1.2.1 La croissance, ses relations avec les variables environnementales et la composition isotopique de carbone

Parmi les différents caractères cibles de la sélection, la croissance est un caractère majeur. La plupart du temps, elle est estimée par des mesures de circonférence et de hauteur, qui permettent d'estimer le volume du tronc et celui à l'hectare. C'est cette grandeur ( $\text{m}^3/\text{ha}$  ou  $\text{m}^3/\text{ha}/\text{an}$ ) qui intéresse directement l'exploitant forestier et que la sélection vise à augmenter.

#### 1.2.1.1 *L'impact des variables environnementales sur la croissance*

La croissance est notamment déterminée par la vitesse de la division cellulaire au niveau des méristèmes primaires (méristème apical) et secondaires (cambium) [130, 131]. Dans le cas des eucalyptus tropicaux, la croissance est continue, que ce soit au niveau du cambium ou des méristèmes apicaux. En effet, contrairement à la plupart des arbres forestiers, ils ne présentent pas d'endodormance. On n'observe donc pas de cernes annuels, contrairement à des espèces de climat tempéré qui présentent un arrêt de croissance pendant l'hiver. Cette croissance continue permet d'obtenir des niveaux de productivité élevés (jusqu'à  $40 \text{ m}^3/\text{ha}/\text{an}$ ) en conditions tropicales pour des hybrides entre *E. urophylla* et *E. grandis* [105]. Ces hauts niveaux de croissance sont souvent associés à des prélèvements de ressources importants [69, 132-134] et à une activité photosynthétique élevée [69]. Par exemple, pour une plantation d'eucalyptus au Brésil, la consommation d'eau annuelle par les arbres (transpiration) a été évaluée à 1394 mm, soit environ 67% des apports en eau (pluie) [135]. Dans le même type de plantation, la consommation en nutriments lors de la première année a été estimée à : 155, 10, 52, 55 et 23 kg/ha de N, P, K, Ca et Mg [134].



	Sensibilité au stress hydrique		
	très sensible	$\Psi_F$ (bar)	peu sensible
	0	10	20
Croissance cellulaire	-----	-----	-----
Synthèse des parois	-----	-----	-----
Synthèse des protéines	-----	-----	-----
Formation des chlorophylles	-----	-----	-----
Niveau de nitrate réductase	-----	-----	-----
Accumulation de AAB	-----	-----	-----
Niveau de cytokinine	-----	-----	-----
Ouverture stomatique	-----	-----	-----
Assimilation du CO <sub>2</sub>	-----	-----	-----
Respiration	-----	-----	-----
Accumulation de proline	-----	-----	-----
Accumulation de sucres	-----	-----	-----

Figure 1.10 : Sensibilité de différents paramètres physiologiques à la diminution du potentiel hydrique foliaire (d'après [136]).

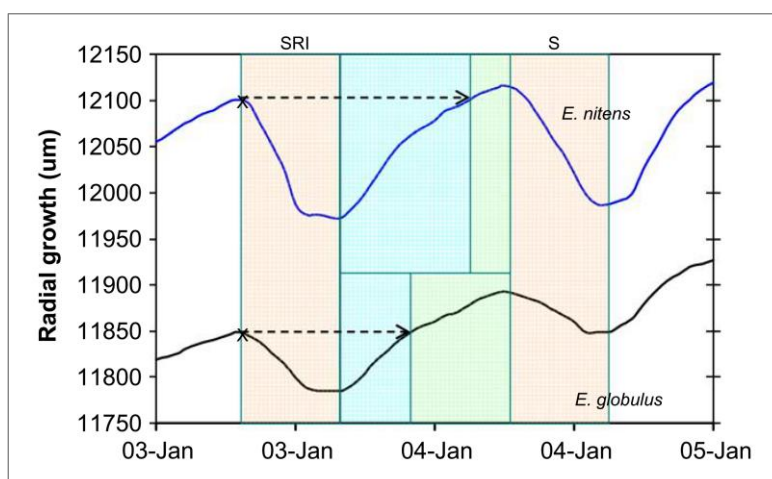


Figure 1.11 : Micro-variations de la croissance radiale pour deux espèces d'*Eucalyptus*. L'évolution temporelle est différente entre les deux espèces. Les auteurs définissent trois phases pour caractériser ces différences : une phase de contraction (S), une phase de reprise (R) et une phase d'évolution nette du rayon (I). D'après [137]





Comme évoqué dans le paragraphe 1.1.2, les processus physiologiques (photosynthèse, absorption de nutriments, transport d'assimilats,...) qui permettent cette croissance dépendent des variables environnementales, et notamment de la disponibilité en eau [64]. La réduction de la croissance fait partie des premières réponses phénotypiques observées suite à l'apparition d'un déficit hydrique (Figure 1.10). La disponibilité en eau et sa variabilité temporelle est donc l'une des principales limitations de la croissance. Des expérimentations au champ sur eucalyptus ont montré un gain de croissance de 30% au Brésil [138] et 40% en République du Congo [139] en levant cette contrainte grâce à l'irrigation. Cette différence de productivité souligne l'importance de la saison sèche notamment dans le cadre des plantations d'eucalyptus en République du Congo.

### 1.2.1.2 *La dynamique de croissance*

Bien que la mesure que l'on en fait soit la plupart du temps ponctuelle, la croissance est un processus dynamique. De ce fait, de nombreuses études ont cherché à modéliser la trajectoire de croissance en fonction du temps et de l'environnement [140-143]. Ces modèles sont utilisés pour des échelles de temps longues (durée de rotation), afin d'améliorer la gestion et les pratiques sylvicoles, et de prédire la production de biomasse. Les premiers modèles utilisés sont des modèles empiriques basés sur des mesures du diamètre à différents âges [142]. Avec l'amélioration de la compréhension du fonctionnement physiologique de l'arbre et de son interaction avec les facteurs climatiques, des modèles mécanistes ont ensuite été élaborés. Ils tiennent principalement compte de l'allocation du carbone, de l'interception de la lumière, de la disponibilité en eau et en nutriments, des effets du climat et de la sylviculture [74, 144, 145]. Ces modèles nécessitent de nombreuses mesures pour être calibrés puis validés, et sont souvent adaptés à un type de plantation donné.

À une échelle beaucoup plus fine, la croissance est également étudiée à travers les variations journalières du rayon du tronc (Figure 1.11). L'acquisition de mesures à cette échelle se fait directement au niveau du tronc grâce à des micro-dendromètres. Ces variations, qui sont de l'ordre de la centaine de  $\mu\text{m}$ , ont été mises en relation avec le statut hydrique de l'arbre et les flux d'eau dans le tronc [146-152]. Il existe une latence entre la variation du potentiel hydrique foliaire (ou des branches) et les variations du rayon du tronc. Cette latence est notamment expliquée par la capacité de stockage (capacitance) du tronc [147, 151].

**Encadré 3 : Les isotopes stables du carbone et l'efficacité d'utilisation de l'eau**

La composition en isotope stable du carbone ( $^{12}\text{C}$  et  $^{13}\text{C}$ ) d'un échantillon ( $\delta^{13}\text{C}$ ) est définie de manière relative par rapport à un standard qui est en général la Pee Dee belemnite (PDB, [153]) et est exprimée en pour mille (‰) :

$$\delta^{13}\text{C} = \left( \frac{{}^{13}\text{C}/{}^{12}\text{C}_{\text{éch}}}{{}^{13}\text{C}/{}^{12}\text{C}_{\text{PDB}}} - 1 \right) \times 1000$$

Du fait de la différence de taille entre le  $^{12}\text{CO}_2$  et le  $^{13}\text{CO}_2$ , il existe une discrimination par les plantes en faveur de la molécule la plus petite ( $^{12}\text{CO}_2$ ). Cette discrimination isotopique du carbone ( $\Delta$ ) est définie par l'équation suivante [154]:

$$\Delta = \frac{\delta_{\text{air}} - \delta_{\text{plante}}}{1 + \delta_{\text{plante}}/1000}$$

avec une composition isotopique de l'air ( $\delta_{\text{air}}$ ) d'environ -8‰. Pour les plantes en C3 (la majorité des arbres), deux étapes principales interviennent dans la discrimination à l'égard  $^{13}\text{C}$  : (a) la diffusion du  $\text{CO}_2$  à travers les stomates et (b) la carboxylation du ribulose biphosphate par la RuBisCO [155-157]. D'autres processus interviennent dans la formation de  $\Delta$  mais sont en général négligés: la dissolution et diffusion du  $\text{CO}_2$  dans la phase liquide et la photorespiration [155-157]. L'équation résultante est la suivante [156]:

$$\Delta = a + (a - b) \frac{C_i}{C_a}$$

avec  $C_a$  la concentration en  $\text{CO}_2$  de l'atmosphère et  $C_i$  celle de l'espace intercellulaire. Farquhar *et al.* [156] ont mis en lien  $\Delta$  et l'efficacité d'utilisation de l'eau intrinsèque (water use efficiency,  $\text{WUE}_i$ ) qui est défini comme le rapport entre l'assimilation de  $\text{CO}_2$  ( $A$ ) et la conductance stomatique pour l'eau ( $g_s$ ) [158]. De manière simplifiée, cela donne l'équation suivante :

$$\text{WUE}_i = \frac{A}{g_s} = \frac{C_a}{1.6} \times \frac{b - \Delta}{b - a}$$

$\Delta$  est donc négativement relié à la  $\text{WUE}_i$ . D'abord mise en évidence chez les plantes annuelles [154, 158], cette relation a également été chez les arbres forestiers bien qu'il existe une variabilité importante du  $\delta^{13}\text{C}$  entre le différent compartiment de l'arbre [159, 160].



L'arbre représente l'interface entre le sol et l'atmosphère dans le continuum sol-plante-atmosphère. Son tronc sert de tampon entre les feuilles qui transpirent et les racines qui absorbent de l'eau. Au début de la journée, l'arbre transpire et mobilise l'eau contenue dans le tronc, ce qui provoque la réduction de son rayon. Lorsque la transpiration s'arrête, l'arbre retourne à un équilibre hydrique grâce à l'absorption racinaire, on observe alors une augmentation du rayon. (Figure 1.11) Le signal des micro-variations du rayon est donc un indicateur du contenu en eau du tronc, mais il intègre également la croissance radiale (division du cambium, élongation cellulaire). Selon l'échelle d'analyse, l'information apportée par le signal sur le statut hydrique et la croissance radiale n'est pas la même. Sur un pas de temps journalier, bien que de la croissance ait lieu, le signal traduit majoritairement le fonctionnement hydrique. Sur une échelle de temps plus longue (semaine, mois), il traduit essentiellement la croissance.

#### 1.2.1.3 *La composition isotopique du carbone de l'arbre, un indicateur de l'efficience d'utilisation de l'eau et son lien avec la croissance*

Le lien entre la croissance de l'arbre et l'eau peut également être abordé sous l'angle de l'efficience d'utilisation de l'eau (water use efficiency, WUE). En effet, la WUE est définie à l'échelle de l'arbre comme le rapport entre la quantité de biomasse et la quantité d'eau transpirée pour la produire [132]. La WUE est principalement liée à l'absorption de lumière, à la capacité photosynthétique des feuilles, à la transpiration de la canopée et à la conductance stomatique, qui est en partie déterminée par les caractéristiques hydrauliques des différents compartiments de l'arbre (racines, tronc et feuilles) [69, 161]. À l'échelle de la feuille, la WUE intrinsèque peut être définie comme le rapport entre l'assimilation du  $\text{CO}_2$  ( $A$ ) et la conductance stomatique pour l'eau ( $g_s$ ). Cette grandeur a été mise en relation avec la discrimination par les plantes entre les deux isotopes stables du carbone ( $^{12}\text{C}$  et  $^{13}\text{C}$ ) et par extension avec la composition isotopique du carbone ( $\delta^{13}\text{C}$ ) de la plante (Encadré 3, [154, 156, 158]). De ce fait le  $\delta^{13}\text{C}$  a été utilisé comme un indicateur intégré et robuste de la WUE dans de nombreuses études [158, 162-164].

Chez les arbres forestiers, la relation entre croissance et  $\delta^{13}\text{C}$  (et par extension la WUE) a été étudiée chez de nombreuses espèces, mais aucune tendance claire ne se dégage. En effet, des corrélations phénotypiques positives (*Eucalyptus globulus* [165], *Pinus pinaster* [166, 167], *Pinus elliottii* [168], *Populus deltoides*×*Populus trichocarpa* [169]), nulles (*Araucaria*

**Encadré 4 :** Notions d'interaction génotype-environnement (G×E) et de plasticité phénotypique.

Une modification de l'environnement peut induire ou non une réponse des génotypes considérés. Si cette réponse est différente d'un génotype à un autre il y a interaction. Cette interaction peut être intégrée dans le cadre du modèle de génétique quantitative (Encadré 2) :

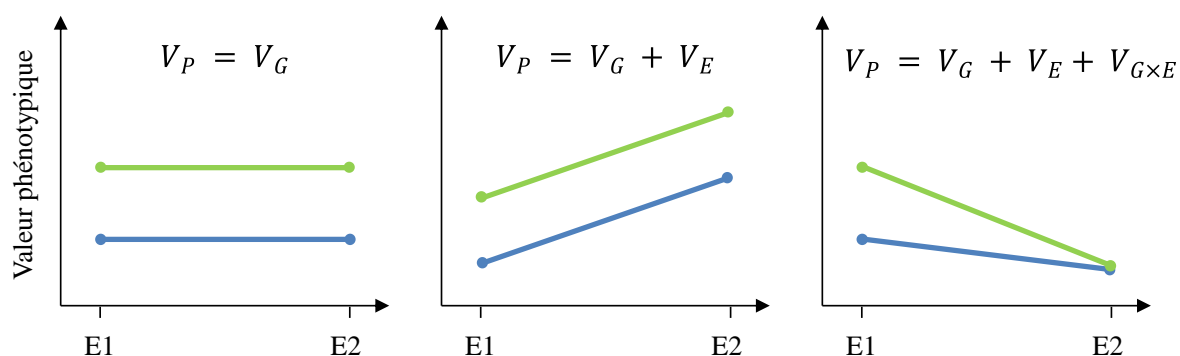
$$P = G + E + G \times E$$

La variance associée est modélisée de la manière suivante :

$$V_P = V_G + V_E + V_{G \times E}$$

Dans ce modèle, la part de variance environnementale ainsi que celle associée à l'interaction G×E représentent la plasticité phénotypique.

La réponse de deux génotypes à un changement d'environnement peut être représentée schématiquement de la manière suivante :



Cette représentation du comportement d'un génotype en fonction de l'environnement est également appelée norme de réaction. L'environnement peut correspondre à une variation du milieu ou à une variation temporelle.



*cunninghamii* [170], *Eucalyptus* [171], *Populus deltoides*×*Populus nigra* [172], *Populus nigra* [173]) ou négatives (*Eucalyptus globulus* [174, 175], *Pinus radiata* [176]) ont été mises en évidence entre ces deux caractères. Cette variabilité entre études souligne les différents mécanismes physiologiques qui conduisent au  $\delta^{13}\text{C}$  et par extension à la WUE, mais également la variabilité présente au sein de l'arbre [177]. Dans le cas d'une corrélation positive, l'assimilation de la feuille serait prépondérante dans l'élaboration de la WUE. À l'inverse, la conductance stomatique serait limitante dans le cas d'une corrélation négative. Cette variabilité peut également, s'expliquer par l'effet marqué des variations environnementales notamment de disponibilité en eau sur  $\delta^{13}\text{C}$  [168, 171, 176] mais également sur la croissance. De plus, plusieurs étapes de discrimination des isotopes du carbone existent entre l'assimilation du carbone au niveau foliaire et la croissance du tronc ce qui complexifie le lien entre  $\delta^{13}\text{C}$  et croissance [159].

## 1.2.2 Les bases génétiques de la croissance

### 1.2.2.1 Un déterminisme génétique relativement faible

Comme le suggèrent les observations au niveau phénotypique, l'héritabilité de la croissance, au sens large comme au sens strict, est en général assez faible (0,1 à 0,4 ; [120, 122, 123, 178, 179]) quelle que soit l'espèce considérée. Même si aucune tendance claire ne se dégage des différentes études, le déterminisme génétique de la croissance varie dans le temps. Selon Franklin [180], d'après des observations sur conifères (*Pinus ponderosa*, *Pinus taeda*, *Pinus elliotii* et *Pseudotsuga menziesii*), trois phases peuvent être observées : une phase juvénile caractérisée par une héritabilité faible (acclimatation), une phase mature avec une héritabilité élevée, et une phase de suppression avec une chute d'héritabilité associée à la mort des arbres les plus faibles. Ce modèle n'a toutefois pas été confirmé sur *Eucalyptus* [181, 182] ou sur *Pinus sylvestris* [183]. Du fait du faible nombre de mesures pour caractériser la croissance dans la majorité des études, l'évolution temporelle de l'héritabilité a été décrite en utilisant des modèles de croissance [184, 185]. Par ailleurs, l'interaction génotype-environnement (G×E, Encadré 4) est un facteur clé dans la compréhension du déterminisme génétique de la croissance.

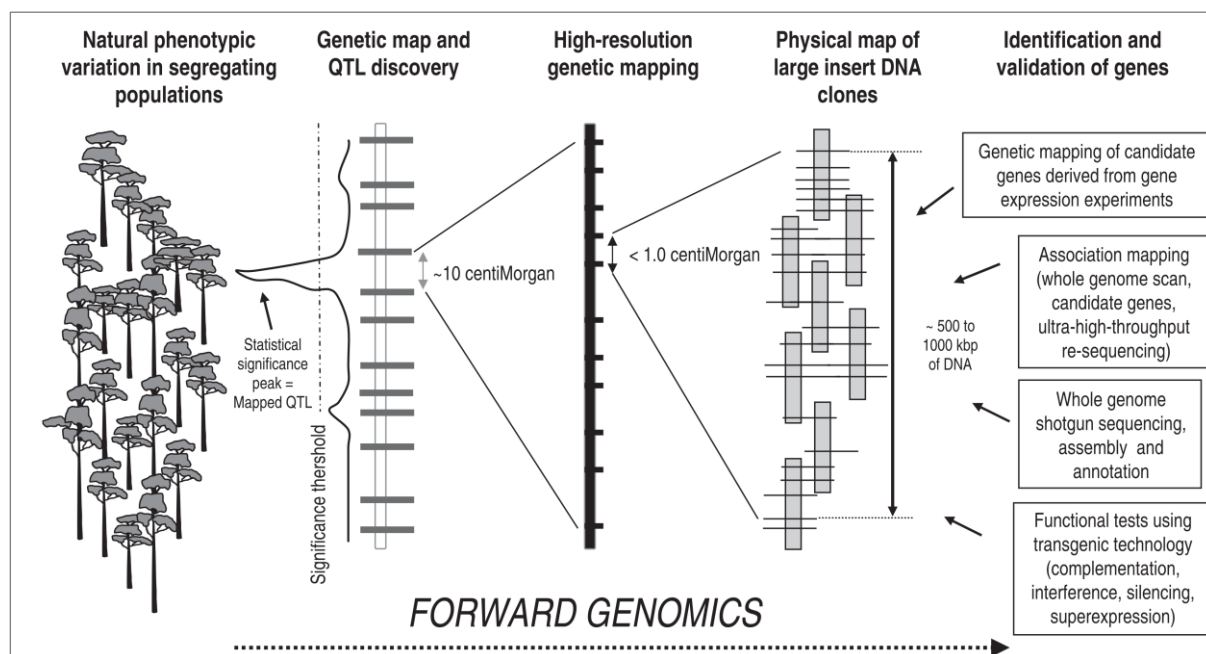


Figure 1.12 : Représentation schématique de l'approche de génomique directe utilisant l'analyse QTL. La cartographie génétique et l'analyse QTL sont utilisées dans des populations en ségrégation, spécifiquement créées pour présenter une variabilité phénotypique importante pour le caractère cible (croissance). Les QTL majeurs détectés sont localisés sur une carte génétique haute résolution qui permet d'affiner la région (cette étape nécessite de grands effectifs). La région génomique d'intérêt est ensuite projetée sur la séquence d'un génome de référence afin d'analyser son contenu en gènes. Plusieurs approches de génomique inverse peuvent alors être utilisées pour identifier, tester et valider les gènes sous-jacents aux QTL ciblés. D'après [126]



À côtés des méthodes basées sur le phénotype, qui ont montré leur efficacité en sélection, les approches intégrant l'information moléculaire sont en train de se développer. La nécessité de mettre en place des approches plus rapides, pouvant intégrer des caractères plus complexes motive cette évolution.

#### 1.2.2.2 L'architecture génétique de la croissance

L'analyse QTL est une approche dite de "génétique directe" qui part de la variabilité du phénotype pour détecter les zones chromosomiques impliquées dans son déterminisme (Figure 1.12). Elle permet d'étudier l'architecture génétique de caractères complexes comme la croissance, c'est-à-dire de déterminer le nombre, la localisation et l'effet des régions génomiques impliquées dans le déterminisme génétique du caractère. Cette approche ainsi que la cartographie génétique (support de l'analyse QTL) sont détaillées dans les parties 2.6 et 2.7 du matériel et méthodes de cette thèse.

De nombreuses études de détection de QTL ont été réalisées chez les arbres forestiers depuis la création des premières cartes génétiques [186, 187]. Des QTL pour les caractères liés à la croissance ont été détectés dans les principales espèces d'arbres utilisées en plantation [124, 188]. Que ce soit pour le diamètre ou la circonférence, la variation phénotypique expliquée par les QTL est généralement faible (de l'ordre de quelques %) mais peut dans certain cas atteindre jusqu'à 20 à 30 % [189-193]. Ces forts pourcentages de variance expliquée sont probablement surestimés, en raison des faibles tailles d'échantillons utilisés dans la plupart des études [194].

À l'instar de l'héritabilité, l'architecture génétique de la croissance varie en fonction du temps et de l'environnement. En effet, différentes études ont montré une variabilité dans la localisation et l'effet des QTL de croissance en fonction de l'environnement : l'interaction QTL-environnement [195-201]. Par exemple sur le saule, Rönnberg-Wästljung et al. [195] ont trouvé 14% de QTL communs entre traitement sec et irrigué pour la hauteur et 28% pour la circonférence. Un pourcentage plus élevé de QTL communs (60%) a été mis en évidence pour la croissance en hauteur entre deux milieux (naturel et contrôlé) sur *Picea glauca* [200]. Pour des essais au champ sur *Eucalyptus globulus*, Freeman et al. [201] ont mis en évidence que 38% des QTL de croissance présentaient de l'interaction G×E. Malgré une variabilité entre études, qui est à relier avec l'espèce et les conditions de croissance plus ou moins stressantes, l'instabilité des QTL de croissance en fonction de l'environnement semble forte.



Une évolution des QTL de croissance au cours du temps a également été rapportée. Emebiri *et al.* [190] ont caractérisé trois types de QTL en fonction de leur évolution temporelle : précoce, intermédiaire et tardif. Cette instabilité temporelle est d'autant plus marquée que l'intervalle de temps entre mesures est long [191, 199, 202, 203]. Toutefois, ces études ne présentent qu'un faible nombre de mesures pour caractériser la croissance (4 à 5 tout au plus), et pour certaines d'entre elles, seule une faible part de la durée de rotation est étudiée. Cette instabilité des QTL de croissance en fonction du temps est à mettre en relation avec l'impact de l'environnement sur la dynamique de croissance, et l'évolution de l'héritabilité de ce caractère au cours du temps.

Un dernier facteur d'instabilité des QTL de croissance est le fond génétique. En effet, les études qui ont conduit des analyses QTL sur plusieurs croisements montrent que les régions mises en évidence ne sont pas forcément répétables [204]. Pour *Picea glauca*, seulement 20% des QTL mis en évidence étaient communs entre deux familles de plein-frères non apparentées [200]. La même tendance a été observée chez *Eucalyptus globulus*, avec seulement 16% de QTL communs entre 4 familles de plein-frères [201]. Cette instabilité peut s'expliquer par des effets alléliques différents entre familles pour les différents QTL, avec une possibilité d'interaction G×E. Une autre cause de l'instabilité peut être le manque de marqueurs communs entre croisements, ce qui limite la colinéarité entre chaque carte parentale et donc limite la possibilité d'identifier des QTL communs ou spécifiques.

Ces différentes contraintes à l'identification de régions génomiques stables pour la croissance ont limité cette approche au domaine de la recherche. Néanmoins l'arrivée des nouvelles technologies de séquençage et de génotypage à haut débit devrait permettre de surmonter certaines de ces limites. Tout d'abord, les marqueurs développés récemment sont plus facilement transférables entre pedigrees ou espèces [205-208]. Pour l'*Eucalyptus*, une puce DArT a été développée [205], permettant la cartographie génétique comparée entre plusieurs espèces [209], mais aussi l'analyse comparative des QTL pour les caractères de croissance entre différentes études. De plus, la disponibilité de génome annoté comme celui du *P. trichocarpa* [210] ou d'*E. grandis* (<http://www.phytozome.net/eucalyptus.php>), et celui de *Picea abies* [211] va permettre d'identifier des gènes candidats dans les intervalles de confiance des QTL. Cette démarche sera facilitée par l'obtention de cartes génétiques précises, grâce au génotypage d'un grand nombre d'individus et de marqueurs. Car





contrairement aux espèces de grandes cultures, il n'existe pas de population hautement recombinante chez les arbres. La cartographie génétique est donc basée sur des F1, F2 ou BC.

La question majeure associée à l'étude de l'architecture génétique est donc l'effet de l'environnement sur la croissance, et plus particulièrement l'étendue de l'interaction  $G \times E$ . Dans un contexte de variabilité des conditions environnementales (disponibilité en eau) lié au changement climatique, la prise en compte de l'interaction  $G \times E$  est essentielle dans la caractérisation de l'architecture génétique de caractères complexes comme la croissance.

### 1.2.3 Stratégie d'intégration de l'interaction $G \times E$ pour l'étude de l'architecture génétique de la croissance

Il existe plusieurs approches pour intégrer l'interaction  $G \times E$  dans l'analyse de l'architecture des caractères complexes tels que la croissance [212-215]. Elles peuvent être séparées en deux catégories. La première comprend des approches centrées sur le génotype (marqueurs moléculaires), reposant sur l'analyse comparative des QTL détectés dans plusieurs environnements. Pour ces approches, l'interaction QTL x environnement ( $QTL \times E$ ) peut être testée de manière statistique. La seconde rassemble différentes approches basées sur une modélisation du phénotype, pour obtenir des variables synthétiques qui caractérisent le génotype indépendamment de l'environnement. Dans ces différentes approches l'environnement peut alors traduire une variation des conditions du milieu (sites d'étude différents, modification d'une ou plusieurs variables environnementales en milieu contrôlé) ou bien une évolution temporelle (trajectoire de croissance).

Cette partie donne l'éventail des possibilités pour étudier l'interaction  $QTL \times E$  mais ne se veut pas exhaustive au niveau des modèles statistiques utilisés car ils sont nombreux et pour certains spécifiques à l'espèce ou à l'expérimentation (cas de la modélisation du phénotype). Deux revues de van Eeuwijk *et al.* [212, 213] fournissent l'écriture formelle et des applications sur un large éventail de modèles.

#### 1.2.3.1 *Les approches centrées sur le génotype*

La première méthode parmi les approches centrées sur le génotype (QTL) est la comparaison, a posteriori, de la localisation et des effets des QTL détectés dans plusieurs environnements. L'interaction  $QTL \times E$  n'est alors pas modélisée de manière statistique, c'est



la présence ou l'absence de QTL dans les différents environnements qui nous renseigne indirectement sur l'interaction QTL×E. Deux types de QTL peuvent alors être distingués :

- Des QTL détectés dans tous les environnements. Il n'y a donc pas d'interaction QTL×E, ce sont donc des QTL constitutifs car ils sont indépendants des conditions environnementales.
- Des QTL détectés seulement dans certains environnements. Il y a donc une interaction QTL×E possible.

Cette approche qualitative a été appliquée dans différentes études chez les arbres forestiers, pour caractériser la stabilité de l'architecture génétique de la croissance [189, 193, 195, 200, 216]. Ces études donnent un aperçu de l'effet de l'environnement sur l'architecture génétique de la croissance, mais ne permettent pas de quantifier le niveau d'interaction.

D'autres approches ont été développées pour tester de manière plus formelle l'interaction QTL×E grâce à des modèles statistiques. La méthode la plus simple consiste en une analyse de variance à partir des QTL obtenus dans différents environnements. Elle a notamment été employée dans une étude sur *Pseudotsuga menziesii* [217] et une plus récente sur *E. globulus* [201]. L'avantage de cette méthode est sa simplicité de mise en œuvre, mais elle ne met pas à profit les répétitions fournies par les différents environnements. Afin d'utiliser l'ensemble de l'information fournie par les différents environnements, de nombreuses méthodes d'analyse QTL ont été proposées [218-226]. En plus de tester l'hypothèse d'interactions QTL×E, le traitement simultané des données provenant de plusieurs environnements peut augmenter significativement la puissance de détection des QTL, avec une meilleure estimation des positions et des effets. Ces différentes méthodes se différencient notamment par le cadre statistique choisi et les limites qu'il implique. Certaines sont basées sur l'utilisation de modèles de mélange [218, 219], d'autres sur des méthodes de régression [220, 222], l'utilisation de modèles mixtes [221, 223-225] ou sur une approche bayésienne [226]. D'abord développée pour les espèces de grandes cultures [227-229], pour lesquelles les expérimentations multi-site sont courantes, ces approches ont également été utilisées chez les arbres forestiers, surtout dans le cadre d'approche pluriannuelle [230] plutôt que multi-site. Le faible nombre d'études utilisant ces méthodes reflète la difficulté de mise en place et de suivi expérimental multi-environnement chez ces espèces pérennes.



### 1.2.3.2 Les approches basées sur la modélisation du phénotype

La prise en compte de l'interaction  $G \times E$  pour l'étude de l'architecture génétique de caractères complexes a également été abordée sous l'angle du phénotype. Ces approches utilisent la modélisation du phénotype afin de minimiser ou d'intégrer l'effet de l'environnement sur le caractère étudié. Elles peuvent être conduites en deux étapes : la modélisation du phénotype puis la détection de QTL conventionnelle, ou en une seule étape dans un même cadre statistique qui combine la modélisation du phénotype et la détection de QTL. Dans les deux cas, les modèles utilisés pour décrire le phénotype sont du même type.

#### La modélisation du phénotype

L'objectif de la modélisation du phénotype est d'obtenir une ou plusieurs variables synthétiques, propres à chaque génotype, à partir de l'ensemble des données phénotypiques obtenues dans différents environnements. Deux types d'approches sont principalement utilisés.

La première approche de modélisation se base sur des modèles linéaires à effets fixes ou mixtes, qui intègrent l'ensemble du plan d'expérience pour modéliser l'interaction  $G \times E$ . Elle comprend de nombreuses méthodes, qui diffèrent par la modélisation de l'interaction  $G \times E$  [231]. Les méthodes les plus courantes sont :

- La décomposition en valeurs singulières, comme les modèles AMMI (additive main effects and multiplicative interaction, [232]) qui modélisent l'interaction  $G \times E$  de façon multiplicative grâce à une analyse en composantes principales.
- La régression factorielle, qui utilise directement des données environnementales pour caractériser l'interaction  $G \times E$ . Un cas particulier est la régression de Finlay et Wilkinson [233].
- Le modèle mixte qui contrairement aux deux premières méthodes, modélise l'interaction  $G \times E$  en effet aléatoire. Cette méthode peut aussi intégrer les deux précédentes.

À partir de ces différentes méthodes on peut donc estimer (ou prédire dans le cadre du modèle mixte : BLUP) des paramètres de l'interaction  $G \times E$  ou des effets génétiques pour chaque génotype. Ces différents paramètres sont ensuite utilisés en analyse QTL classique.

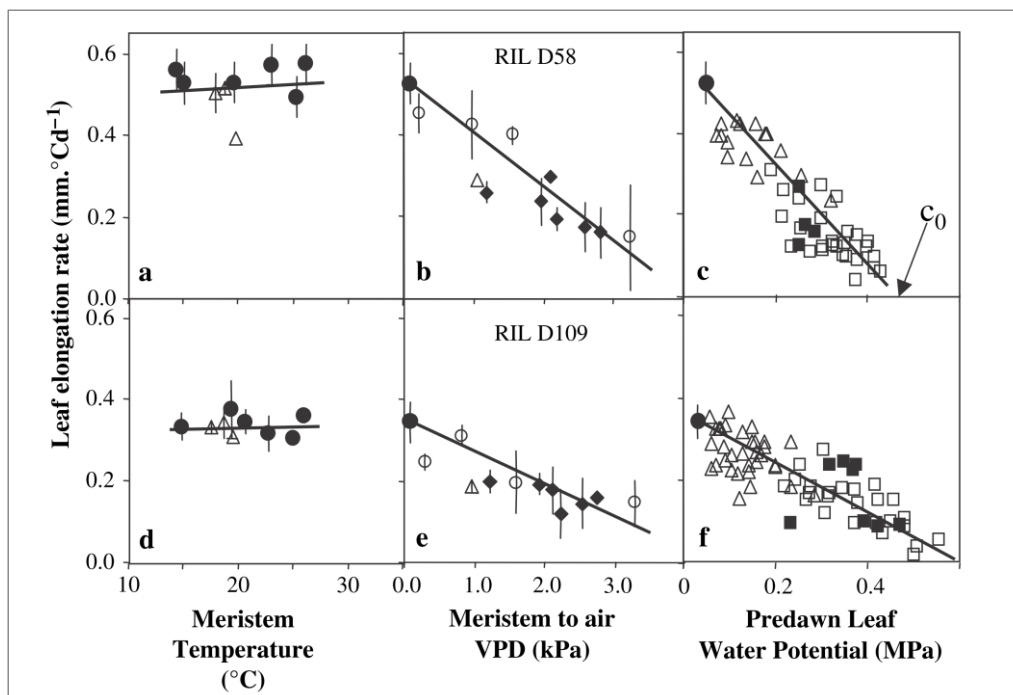


Figure 1.13 : Réponse de l'élongation foliaire chez le maïs à la température (a, d), au déficit de pression de vapeur (b, e) et au contenu en eau du sol (c, f) pour deux génotypes (RILs). Les cercles représentent une expérimentation en serre sans déficit hydrique du sol (symbole plein : nuit, symbole ouvert : jour), les losanges une expérimentation au champ, les triangles une expérimentation en serre avec un déficit hydrique du sol faible, les carrés une expérimentation en serre avec un déficit hydrique du sol modéré. D'après [234].

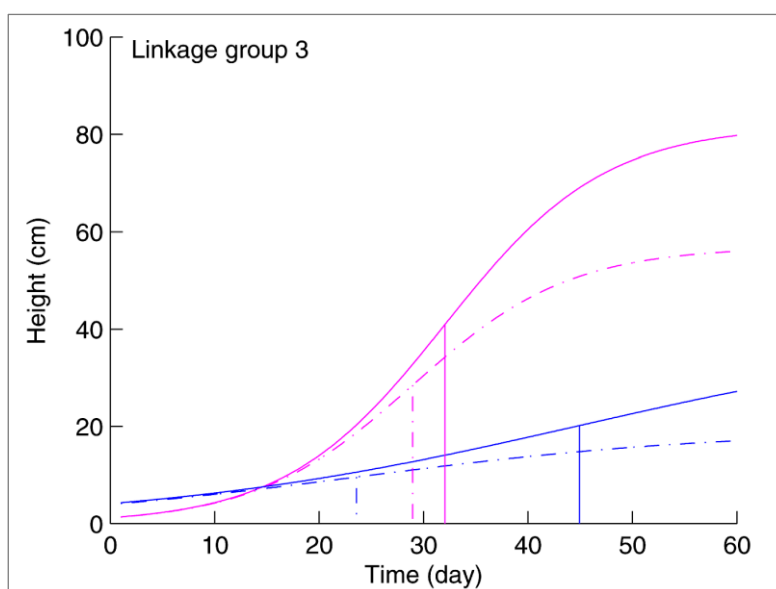


Figure 1.14 : Exemple de résultat d'analyse QTL avec l'approche de « functional mapping » sur la croissance en hauteur du soja. Les courbes de croissance pour les deux classes de génotype (ligne pleine et ligne pointillée) sont représentées au niveau d'un QTL. Les deux années de mesure sont représentées en bleu et en rose. Les barres verticales indiquent le point d'inflexion de la courbe. D'après [235].



Les travaux de Rae *et al.* [198] sur peuplier sont un exemple d'utilisation des modèles AMMI pour étudier l'architecture génétique de l'interaction G×E.

La seconde approche consiste à modéliser le phénotype en fonction d'une variation environnementale (température, luminosité, contenu en eau du sol) ou temporelle pour obtenir des courbes de réponse, des normes de réaction ou des trajectoires de croissance. Cette approche utilise des modèles linéaire, non-linéaire ou non-paramétrique pour ajuster les données phénotypiques. Les courbes obtenues sont définies par un nombre réduit de paramètres qui ont une signification biologique (la sensibilité au déficit hydrique, la vitesse de croissance etc.) et qui caractérisent les génotypes. En général, les courbes de réponse fondées sur des relations empiriques sont établies dans différentes conditions expérimentales (conditions contrôlées, au champ) et sont valables sur une large gamme de conditions environnementales pour un génotype donné (Figure 1.13). Par conséquent, elles peuvent être considérées comme une caractéristique stable d'un génotype. Cette approche a été appliquée dans le cadre d'analyses QTL sur différentes espèces : *Arabidopsis thaliana* [236], *Zea mays* [234, 237], *Hordeum vulgare* [238].

### **Les approches combinées**

Aussi appelées « functional mapping » [239], ces approches font la synthèse entre la modélisation du phénotype et l'analyse QTL. Elles ont été développées plus particulièrement pour l'analyse des courbes de réponse [235, 240-246]. Bien que plus difficiles à implémenter, ces approches apportent plusieurs avantages, comme la prise en compte de l'incertitude sur les paramètres du modèle de croissance dans l'analyse QTL. Ces approches ont été appliquées pour décrire l'architecture génétique de la dynamique de croissance chez le peuplier [239, 247, 248] et le soja (Figure 1.14, [235]), ainsi que la sénescence foliaire chez la pomme de terre [246]. Malgré la puissance de ce type d'approche pour l'étude de la croissance et de sa dynamique, elle reste encore peu utilisée.

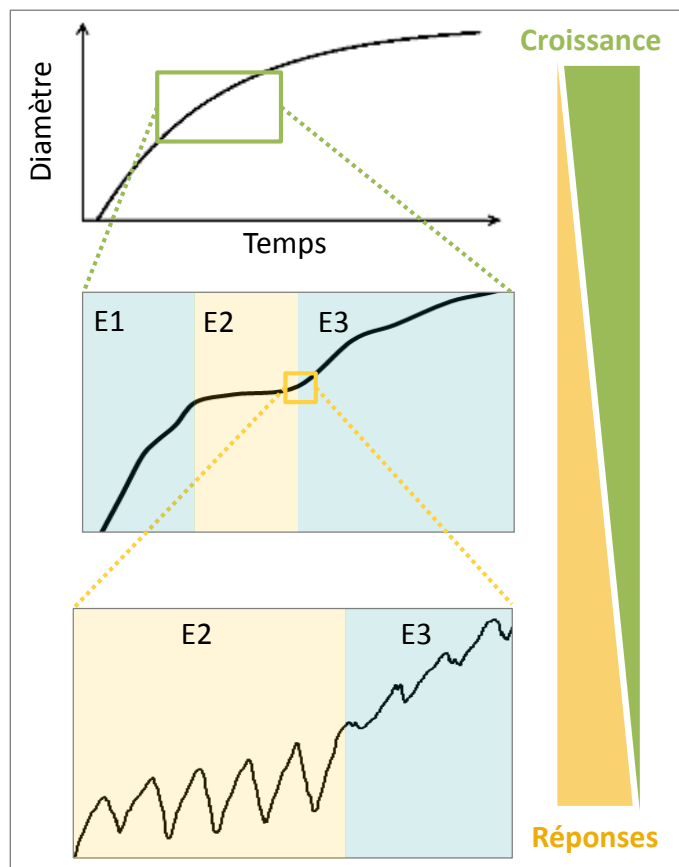


Figure 1.15 : Échelles d'étude de la croissance utilisées au cours de cette thèse. De la trajectoire de croissance globale jusqu'à la caractérisation en continu des micro-variations du rayon.

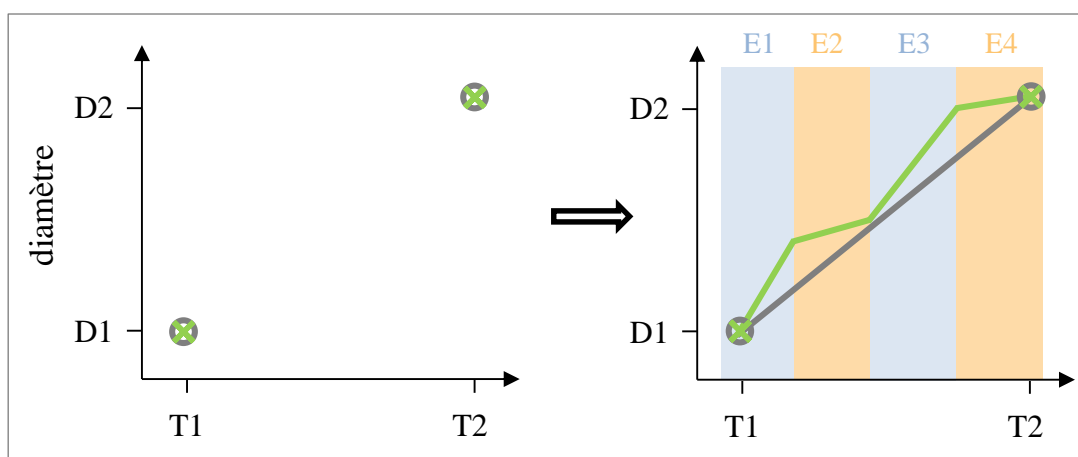


Figure 1.16 : Représentation schématique de la mise en évidence de l'interaction  $G \times E$  grâce à la prise en compte de la trajectoire de croissance à une échelle plus fine. À gauche, les génotypes (gris et vert) ont le même phénotype quel que soit le temps (ou l'environnement) il n'y donc pas d'interaction  $G \times E$ . À droite, les trajectoires de croissance sont différentes entre génotypes en fonction des différents environnements traversés, (E1 à E4) ce qui permet de mettre en évidence une interaction  $G \times E$ .



### 1.3 Objectif de la thèse et démarche adoptée

Dans un contexte d'évolution des contraintes environnementales et d'augmentation de la demande en bois au niveau mondial, le maintien de la productivité des forêts plantées est un enjeu majeur. Une sélection plus efficace, basée sur l'information génomique est une des pistes envisageable pour y répondre. Cependant, plusieurs freins à sa mise en place demeurent. Les ressources génomiques n'étant plus le facteur limitant, grâce à l'accélération des technologies de séquençages et de génotypage haut débit et à la baisse des coûts, le principal défi reste la prise en compte de l'interaction  $G \times E$  dans l'étude de l'architecture génétique d'un caractère. Ceci est d'autant plus marqué pour un caractère complexe comme la croissance, qui varie en fonction des contraintes environnementales mais aussi en fonction du temps. L'analyse des bases génétique, ontologique et environnementale de la croissance est donc au cœur des recherches sur l'amélioration génétique des arbres forestiers. L'eucalyptus, grâce à sa croissance rapide (rotation entre 6 et 10 ans) et à la disponibilité d'un génome de référence, constitue un système biologique unique pour disséquer le déterminisme génétique et environnemental de la dynamique de croissance.

L'objectif de la thèse est de comprendre le déterminisme génétique et environnemental de la croissance, par l'analyse de l'architecture génétique de la dynamique de croissance et de l'efficacité d'utilisation de l'eau chez l'*Eucalyptus*. Pour le premier volet, différentes échelles de temps en relation avec les variations environnementales et notamment la disponibilité en eau ont été étudiées. Pour le second volet, la composition isotopique du carbone a été utilisée comme indicateur de l'efficacité d'utilisation de l'eau. Cet objectif se décompose en plusieurs questions de recherches associées à des hypothèses de travail :

- La prise en compte de la dynamique de croissance permet-elle d'expliquer l'instabilité de l'architecture génétique de la croissance ?

*Hypothèse : La croissance totale (radiale et apicale) est un caractère intégrateur des différents environnements traversés par l'arbre. L'étude de la dynamique de croissance à différentes échelles de temps permet de décomposer les réponses de l'arbre aux variations environnementales (Figure 1.15).*

*Hypothèse : La dynamique de croissance est un indicateur de la plasticité aux contraintes abiotiques (surtout hydrique) qui permet une meilleure prise en compte de l'interaction  $G \times E$  (Figure 1.16).*



- Existe-il un déterminisme génétique commun entre la croissance et le  $\delta^{13}\text{C}$  ?

*Hypothèse : Le déterminisme génétique entre la croissance et  $\delta^{13}\text{C}$  est différent, bien que des corrélations phénotypique et génétique aient été mises en évidence entre ces deux caractères.*

Pour répondre à cet objectif, un croisement biparental interspécifique *E. urophylla* x *E. grandis* a été étudié dans quatre dispositifs expérimentaux : trois dispositifs QTL et un test clonal. Ces dispositifs ont été plantés à Pointe-Noire en République du Congo. Dans cette région, le climat présente des variations de pluviométrie importantes entre la saison sèche et la saison des pluies. Ces variations entre saisons sont moins fortes pour la température et l'humidité de l'air. Ces conditions sont donc favorables pour observer l'effet de la disponibilité en eau sur la croissance.

L'approche utilisée combine : i/ un génotypage de la descendance de plein-frères, ii/ une caractérisation inter et intra-annuelle de la croissance et du  $\delta^{13}\text{C}$ , ainsi qu'un suivi en continu des micro-variations du diamètre, et iii/ une caractérisation en continu des facteurs environnementaux (température, pluviométrie, humidité du sol,...).





## Références

1. Lambrechts C, Wilkie M, Rucevska I: **Vital forest graphics**: Food and Agriculture Organization of the United Nations (FAO); 2009.
2. FAO: **Global forest resources assessment 2005 : progress towards sustainable forest management**. Rome: Food and Agriculture Organization of the United Nations; 2006.
3. Millennium Ecosystem Assessment: **Ecosystems and human well-being: current state and trends**, vol. 1. Washington DC: Island Press; 2005.
4. FAO: **State of the world's forests, 2011**. In. Rome: Food and Agriculture Organization of the United Nations; 2011.
5. FAO: **Global Forest Resources Assessment 2010: Main Report**: Food and Agriculture Organization of the United Nations; 2010.
6. Schmitt CB, Burgess ND, Coad L, Belokurov A, Besançon C, Boisrobert L, Campbell A, Fish L, Gliddon D, Humphries K *et al*: **Global analysis of the protection status of the world's forests**. *Biological Conservation* 2009, **142**(10):2122-2130.
7. Farley KA, Jobbágy EG, Jackson RB: **Effects of afforestation on water yield: a global synthesis with implications for policy**. *Glob Change Biol* 2005, **11**(10):1565-1576.
8. Bonan GB: **Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests**. *Science* 2008, **320**(5882):1444-1449.
9. Scherr SJ, White A, Kaimowitz D: **A new agenda for forest conservation and poverty reduction**. In: *Forest Trends*, . Washington DC; 2004.
10. Williams M: **Deforesting the Earth: from prehistory to global crisis**. Chicago, États-Unis d'Amérique: University of Chicago Press; 2002.
11. FAO: **State of the world's forests, 2009**. Rome: Food and Agriculture Organization of the United Nations; 2009.
12. Rohde R, Muller RA, Jacobsen R, Muller E, Perlmutter S, Rosenfeld A, Wurtele J, Wickham C: **A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011**. *Geoinformatics & Geostatistics: An Overview* 2013, **01**(01).
13. Carle J, Holmgren P: **Wood from planted forests - a global outlook 2005-2030**. *Forest Products Journal* 2008, **58**(12):7.
14. Evans J: **Planted forests: uses, impacts and sustainability**: CABI and FAO; 2009.
15. Foster G, Rahmstorf S: **Global temperature evolution 1979–2010**. *Environmental Research Letters* 2011, **6**(4):044022.
16. IPCC: **IPCC fourth assessment report climate change 2007**. Geneva: Intergovernmental Panel on Climate Change; 2007.
17. Rohde R, Muller RA, Jacobsen R, Perlmutter S, Rosenfeld A, Wurtele J, Curry J, Wickham C, Mosher S: **Berkeley earth temperature averaging process**. *Geoinformatics & Geostatistics: An Overview* 2013, **1**(2).
18. Hansen J, Ruedy R, Sato M, Lo K: **Global surface temperature change**. *Reviews of Geophysics* 2010, **48**(4):RG4004.
19. IPCC: **Working Group I Contribution to the IPCC Fifth Assessment Report, Climate Change 2013: the Physical Science Basis**. Cambridge: Cambridge University Press; in press.
20. UNEP: **The Emissions Gap Report 2013**. Nairobi: United Nations Environment Programme (UNEP); 2013.
21. IPCC: **Climate Change 2007 : the physical science basis : working group I contribution to the Fourth Assessment Report of the IPCC**. Cambridge: Cambridge University Press; 2007.
22. Easterling WE, Aggarwal P.K., Batima P., Brander K.M., Erda L., Howden S.M., Kirilenko A. , Morton J., Soussana J.-F., Schmidhuber J. *et al*: **Food, fibre and forest products**. In: *Climate Change 2007:*



- Impacts, Adaptation and Vulnerability Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Edited by Parry M.L., Canziani O.F., Palutikof J.P., van der Linden P.J., Hanson C.E. Cambridge: Cambridge University Press; 2007: 273-313.
23. Medlyn BE, Barton CVM, Broadmeadow MSJ, Ceulemans R, De Angelis P, Forstreuter M, Freeman M, Jackson SB, Kellomäki S, Laitat E *et al*: **Stomatal conductance of forest species after long-term exposure to elevated CO<sub>2</sub> concentration: a synthesis**. *New Phytol* 2001, **149**(2):247-264.
  24. Körner C: **Plant CO<sub>2</sub> responses: an issue of definition, time and resource supply**. *New Phytol* 2006, **172**(3):393-411.
  25. Norby RJ, Zak DR: **Ecological Lessons from Free-Air CO<sub>2</sub> Enrichment (FACE) Experiments**. *Annual Review of Ecology, Evolution, and Systematics* 2011, **42**(1):181-203.
  26. Ainsworth EA, Long SP: **What have we learned from 15 years of free-air CO<sub>2</sub> enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO<sub>2</sub>**. *New Phytol* 2005, **165**(2):351-372.
  27. Nowak RS, Ellsworth DS, Smith SD: **Functional responses of plants to elevated atmospheric CO<sub>2</sub>—do photosynthetic and productivity data from FACE experiments support early predictions?** *New Phytol* 2004, **162**(2):253-280.
  28. Sigurdsson BD, Medhurst JL, Wallin G, Eggertsson O, Linder S: **Growth of mature boreal Norway spruce was not affected by elevated [CO<sub>2</sub>] and/or air temperature unless nutrient availability was improved**. *Tree Physiology* 2013, **33**(11):1192-1205.
  29. Dawes M, Hagedorn F, Handa I, Streit K, Ekblad A, Rixen C, Körner C, Hättenschwiler S: **An alpine treeline in a carbon dioxide-rich world: synthesis of a nine-year free-air carbon dioxide enrichment study**. *Oecologia* 2013, **171**(3):623-637.
  30. Körner C, Asshoff R, Bignucolo O, Hättenschwiler S, Keel SG, Peláez-Riedl S, Pepin S, Siegwolf RTW, Zotz G: **Carbon Flux and Growth in Mature Deciduous Forest Trees Exposed to Elevated CO<sub>2</sub>**. *Science* 2005, **309**(5739):1360-1362.
  31. Bader MKF, Leuzinger S, Keel SG, Siegwolf RTW, Hagedorn F, Schleppi P, Körner C: **Central European hardwood trees in a high-CO<sub>2</sub> future: synthesis of an 8-year forest canopy CO<sub>2</sub> enrichment project**. *Journal of Ecology* 2013, **101**(6):1509-1519.
  32. Norby RJ, DeLucia EH, Gielen B, Calfapietra C, Giardina CP, King JS, Ledford J, McCarthy HR, Moore DJP, Ceulemans R *et al*: **Forest response to elevated CO<sub>2</sub> is conserved across a broad range of productivity**. *P Natl Acad Sci USA* 2005, **102**(50):18052-18056.
  33. Moore DJP, Aref S, Ho RM, Pippen JS, Hamilton JG, De Lucia EH: **Annual basal area increment and growth duration of Pinus taeda in response to eight years of free-air carbon dioxide enrichment**. *Glob Change Biol* 2006, **12**(8):1367-1377.
  34. Kostianen K: **Wood properties of northern forest trees grown under elevated CO<sub>2</sub>, O<sub>3</sub> and temperature** PhD. Helsinki: University of Oulu; 2007.
  35. Kaakinen S, Kostianen K, Ek F, Saranpää P, Kubiske ME, Sober J, Karnosky DF, Vapaavuori E: **Stem wood properties of Populus tremuloides, Betula papyrifera and Acer saccharum saplings after 3 years of treatments to elevated carbon dioxide and ozone**. *Glob Change Biol* 2004, **10**(9):1513-1525.
  36. Kilpeläinen A, Peltola H, Ryyppö A, Kellomäki S: **Scots pine responses to elevated temperature and carbon dioxide concentration: growth and wood properties**. *Tree Physiology* 2005, **25**(1):75-83.
  37. Cernusak LA, Winter K, Dalling JW, Holtum JAM, Jaramillo C, Körner C, Leakey ADB, Norby RJ, Poulter B, Turner BL *et al*: **Tropical forest responses to increasing atmospheric CO<sub>2</sub>: current knowledge and opportunities for future research**. *Functional Plant Biology* 2013, **40**(6):531-551.
  38. Körner C: **Responses of Humid Tropical Trees to Rising CO<sub>2</sub>**. *Annual Review of Ecology, Evolution, and Systematics* 2009, **40**(1):61-79.
  39. Medlyn BE, Duursma RA, Zeppel MJB: **Forest productivity under climate change: a checklist for evaluating model studies**. *Wiley Interdisciplinary Reviews: Climate Change* 2011, **2**(3):332-355.



40. Almeida A, Sands P, Bruce J, Siggins A, Leriche A, Battaglia M, Batista T: **Use of a spatial process-based model to quantify forest plantation productivity and water use efficiency under climate change scenarios.** In: *18th World IMACS/MODSIM congress: 2009; Cairns, Australia; 2009*: 13-17.
41. **EucFACE Experiment**
42. Kirschbaum MU: **Does enhanced photosynthesis enhance growth? Lessons learned from CO<sub>2</sub> enrichment studies.** *Plant Physiol* 2011, **155**(1):117-124.
43. Peñuelas J, Canadell JG, Ogaya R: **Increased water-use efficiency during the 20th century did not translate into enhanced tree growth.** *Global Ecology and Biogeography* 2011, **20**(4):597-608.
44. Long SP, Ainsworth EA, Rogers A, Ort DR: **Rising atmospheric carbon dioxide: plants FACE the future.** *Annu Rev Plant Biol* 2004, **55**:591-628.
45. Ainsworth EA, Rogers A, Nelson R, Long SP: **Testing the “source–sink” hypothesis of down-regulation of photosynthesis in elevated [CO<sub>2</sub>] in the field with single gene substitutions in Glycine max.** *Agricultural and Forest Meteorology* 2004, **122**(1–2):85-94.
46. Davey PA, Olcer H, Zakhleniuk O, Bernacchi CJ, Calfapietra C, Long SP, Raines CA: **Can fast-growing plantation trees escape biochemical down-regulation of photosynthesis when grown throughout their complete production cycle in the open air under elevated carbon dioxide?** *Plant, Cell & Environment* 2006, **29**(7):1235-1244.
47. Zotz G, Pepin S, Körner C: **No Down-Regulation of Leaf Photosynthesis in Mature Forest Trees after Three Years of Exposure to Elevated CO<sub>2</sub>.** *Plant Biol (Stuttg)* 2005, **7**(04):369-374.
48. Way DA, Oren R: **Differential responses to changes in growth temperature between trees from different functional groups and biomes: a review and synthesis of data.** *Tree Physiology* 2010, **30**(6):669-688.
49. Kirschbaum MUF: **Direct and Indirect Climate Change Effects on Photosynthesis and Transpiration.** *Plant Biology* 2004, **6**(3):242-253.
50. Sage RF, Kubien DS: **The temperature response of C<sub>3</sub> and C<sub>4</sub> photosynthesis.** *Plant, Cell & Environment* 2007, **30**(9):1086-1106.
51. Morin X, Roy J, Sonié L, Chuine I: **Changes in leaf phenology of three European oak species in response to experimental climate change.** *New Phytol* 2010, **186**(4):900-910.
52. Morin X, Lechowicz MJ, Augspurger C, O’Keefe J, Viner D, Chuine I: **Leaf phenology in 22 North American tree species during the 21st century.** *Glob Change Biol* 2009, **15**(4):961-975.
53. Rossi S, Morin H, Deslauriers A, Plourde P-Y: **Predicting xylem phenology in black spruce under climate warming.** *Glob Change Biol* 2011, **17**(1):614-625.
54. Rossi S, Anfodillo T, Čufar K, Cuny HE, Deslauriers A, Fonti P, Frank D, Gričar J, Gruber A, King GM *et al*: **A meta-analysis of cambium phenology and growth: linear and non-linear patterns in conifers of the northern hemisphere.** *Annals of Botany* 2013, **112**(9):1911-1920.
55. Luo Y, Sherry R, Zhou X, Wan S: **Terrestrial carbon-cycle feedback to climate warming: experimental evidence on plant regulation and impacts of biofuel feedstock harvest.** *GCB Bioenergy* 2009, **1**(1):62-74.
56. Chung H, Muraoka H, Nakamura M, Han S, Muller O, Son Y: **Experimental warming studies on tree species and forest ecosystems: a literature review.** *J Plant Res* 2013, **126**(4):447-460.
57. Wu Z, Dijkstra P, Koch GW, Peñuelas J, Hungate BA: **Responses of terrestrial ecosystems to temperature and precipitation change: a meta-analysis of experimental manipulation.** *Glob Change Biol* 2011, **17**(2):927-942.
58. Bronson DR, Gower ST, Tanner M, Van Herk I: **Effect of ecosystem warming on boreal black spruce bud burst and shoot growth.** *Glob Change Biol* 2009, **15**(6):1534-1543.
59. Wilkening M, Juday GP, Barber VA, Zald HSJ: **Recent climate warming forces contrasting growth responses of white spruce at treeline in Alaska through temperature thresholds.** *Glob Change Biol* 2004, **10**(10):1724-1736.



60. Subedi N, Sharma M: **Climate-diameter growth relationships of black spruce and jack pine trees in boreal Ontario, Canada.** *Glob Change Biol* 2013, **19**(2):505-516.
61. Clark DA, Piper SC, Keeling CD, Clark DB: **Tropical rain forest tree growth and atmospheric carbon dynamics linked to interannual temperature variation during 1984–2000.** *Proceedings of the National Academy of Sciences* 2003, **100**(10):5852-5857.
62. Feeley KJ, Joseph Wright S, Nur Supardi MN, Kassim AR, Davies SJ: **Decelerating growth in tropical forest trees.** *Ecology Letters* 2007, **10**(6):461-469.
63. Wolkovich EM, Cook BI, Allen JM, Crimmins TM, Betancourt JL, Travers SE, Pau S, Regetz J, Davies TJ, Kraft NJB *et al*: **Warming experiments underpredict plant phenological responses to climate change.** *Nature* 2012, **485**(7399):494-497.
64. Pallardy SG: **Physiology of Woody Plants.** In., Third edition edn. San Diego: Academic Press; 2008: 454.
65. Philip JR: **Plant Water Relations: Some Physical Aspects.** *Annual Review of Plant Physiology* 1966, **17**(1):245-268.
66. Tyree MT, Zimmermann MH: **Xylem structure and the ascent of sap,** Second edition edn. Berlin: Springer; 2002.
67. Pereira JS, Pallardy S: **Water Stress Limitations to Tree Productivity.** In: *Biomass Production by Fast-Growing Trees.* Edited by Pereira JS, Landsberg JJ, vol. 166: Springer Netherlands; 1989: 37-56.
68. Kozlowski TT, Kramer PJ, Pallardy SG: **Water Stress.** In: *The Physiological Ecology of Woody Plants.* Edited by Kozlowski TT, Kramer PJ, Pallardy SG. San Diego: Academic Press; 1991: 247-302.
69. Whitehead D, Beadle CL: **Physiological regulation of productivity and water use in Eucalyptus: a review.** *Forest Ecology and Management* 2004, **193**(1–2):113-140.
70. Allen CD, Macalady AK, Chenchouni H, Bachelet D, McDowell N, Vennetier M, Kitzberger T, Rigling A, Breshears DD, Hogg EH *et al*: **A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests.** *Forest Ecology and Management* 2010, **259**(4):660-684.
71. Almeida AC, Soares JV, Landsberg JJ, Rezende GD: **Growth and water balance of Eucalyptus grandis hybrid plantations in Brazil during a rotation for pulp production.** *Forest Ecology and Management* 2007, **251**(1–2):10-21.
72. Wagner F, Rossi V, Stahl C, Bonal D, Hérault B: **Water Availability Is the Main Climate Driver of Neotropical Tree Growth.** *PLoS ONE* 2012, **7**(4):e34074.
73. Stape JL, Binkley D, Ryan MG: **Eucalyptus production and the supply, use and efficiency of use of water, light and nitrogen across a geographic gradient in Brazil.** *Forest Ecology and Management* 2004, **193**(1–2):17-31.
74. Almeida AC, Landsberg JJ, Sands PJ, Ambrogi MS, Fonseca S, Barddal SM, Bertolucci FL: **Needs and opportunities for using a process-based productivity model as a practical tool in Eucalyptus plantations.** *Forest Ecology and Management* 2004, **193**(1–2):167-177.
75. Binkley D, O'Connell AM, Sankaran KV: **Stand development and productivity:** Australian Centre for International Agricultural Research (ACIAR); 1997.
76. Littell JS, Peterson DL, Tjoelker M: **Douglas-fir growth in mountain ecosystems: water limits tree growth from stand to region.** *Ecological Monographs* 2008, **78**(3):349-368.
77. Eilmann B, Rigling A: **Tree-growth analyses to estimate tree species' drought tolerance.** *Tree Physiology* 2012, **32**(2):178-187.
78. Phillips OL, van der Heijden G, Lewis SL, Lopez-Gonzalez G, Aragao LE, Lloyd J, Malhi Y, Monteagudo A, Almeida S, Davila EA *et al*: **Drought-mortality relationships for tropical forests.** *New Phytol* 2010, **187**(3):631-646.
79. Tyree MT, Sperry JS: **Do Woody Plants Operate Near the Point of Catastrophic Xylem Dysfunction Caused by Dynamic Water Stress?: Answers from a Model.** *Plant Physiol* 1988, **88**(3):574-580.



80. Choat B, Jansen S, Brodribb TJ, Cochard H, Delzon S, Bhaskar R, Bucci SJ, Feild TS, Gleason SM, Hacke UG *et al*: **Global convergence in the vulnerability of forests to drought.** *Nature* 2012, **491**(7426):752-755.
81. Wimmer R, Downes GM, Evans R: **Temporal variation of microfibril angle in Eucalyptus nitens grown in different irrigation regimes.** *Tree Physiology* 2002, **22**(7):449-457.
82. Searson MJ, Thomas DS, Montagu KD, Conroy JP: **Wood density and anatomy of water-limited eucalypts.** *Tree Physiology* 2004, **24**(11):1295-1302.
83. Wimmer R, Downes GM, Evans R: **High-resolution analysis of radial growth and wood density in Eucalyptus nitens, grown under different irrigation regimes.** *Ann For Sci* 2002, **59**(5-6):519-524.
84. Flannigan MD, Stocks BJ, Wotton BM: **Climate change and forest fires.** *Sci Total Environ* 2000, **262**(3):221-229.
85. Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW: **Warming and Earlier Spring Increase Western U.S. Forest Wildfire Activity.** *Science* 2006, **313**(5789):940-943.
86. Kirilenko AP, Sedjo RA: **Climate change impacts on forestry.** *Proceedings of the National Academy of Sciences* 2007, **104**(50):19697-19702.
87. Flannigan MD, Logan KA, Amiro BD, Skinner WR, Stocks BJ: **Future Area Burned in Canada.** *Climatic Change* 2005, **72**(1-2):1-16.
88. Williams J, Albright D, Hoffmann AA, Eritsov A, Leonard M: **Findings and implications from a coarse-scale global assessment of recent selected mega-fires.** In: *Fifth International Wildland Fire Conference, Sun City, South Africa: 2011*; 2011.
89. Whittaker JB: **Insects and plants in a changing atmosphere.** *Journal of Ecology* 2001, **89**(4):507-518.
90. Logan JA, Régnière J, Powell JA: **Assessing the impacts of global warming on forest pest dynamics.** *Frontiers in Ecology and the Environment* 2003, **1**(3):130-137.
91. Régnière J: **Prédire la répartition continentale des insectes à partir de leur physiologie.** *Unasylva* 2009, **60**.
92. Carroll AL, Taylor SW, Régnière J, Safranyik L: **Effect of climate change on range expansion by the mountain pine beetle in British Columbia.** In: *Pages 223-232 in TL Shore et al(eds) Mountain Pine Beetle Symposium: Challenges and Solutions, Oct 30-31, 2003 Kelowna BC Natural Resources Canada, Information Report BC-X-399, Victoria: 2004*; 2004.
93. Pinkard EA, Battaglia M, Roxburgh S, O'Grady AP: **Estimating forest net primary production under changing climate: adding pests into the equation.** *Tree Physiology* 2011, **31**(7):686-699.
94. Crozier L, Dwyer G: **Combining population-dynamic and ecophysiological models to predict climate-induced insect range shifts.** *Am Nat* 2006, **167**(6):853-866.
95. Bernier P, Schoene D: **Adapter les forêts et leur gestion aux changements climatiques: un aperçu.** *Unasylva* 2009, **60**.
96. Millar CI, Stephenson NL, Stephens SL: **Climate change and forests of the future: managing in the face of uncertainty.** *Ecological Applications* 2007, **17**(8):2145-2151.
97. Gonçalves JLdM, Alvares CA, Higa AR, Silva LD, Alfenas AC, Stahl J, Ferraz SFdB, Lima WdP, Brancalion PHS, Hubner A *et al*: **Integrating genetic and silvicultural strategies to minimize abiotic and biotic constraints in Brazilian eucalypt plantations.** *Forest Ecology and Management* 2013, **301**(0):6-27.
98. Laclau J-P, Gonçalves JLdM, Stape JL: **Perspectives for the management of eucalypt plantations under biotic and abiotic stresses.** *Forest Ecology and Management* 2013, **301**(0):1-5.
99. Brown AE, Zhang L, McMahon TA, Western AW, Vertessy RA: **A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation.** *Journal of Hydrology* 2005, **310**(1-4):28-61.
100. Scott DF, Prinsloo FW: **Longer-term effects of pine and eucalypt plantations on streamflow.** *Water Resources Research* 2008, **44**(7):W00A08.





101. Collins BM, Stephens SL: **Managing natural wildfires in Sierra Nevada wilderness areas.** *Frontiers in Ecology and the Environment* 2007, **5**(10):523-527.
102. Gonçalves JLM, Stape JL, Laclau JP, Bouillet JP, Ranger J: **Assessing the effects of early silvicultural management on long-term site productivity of fast-growing eucalypt plantations: the Brazilian experience.** *Southern Forests: a Journal of Forest Science* 2008, **70**(2):105-118.
103. du Toit B, Dovey SB, Smith CW: **Effects of slash and site management treatments on soil properties, nutrition and growth of a Eucalyptus grandis plantation in South Africa:** Center for International Forestry Research (CIFOR); 2008.
104. du Toit B, Smith CW, Little KM, Boreham G, Pallett RN: **Intensive, site-specific silviculture: Manipulating resource availability at establishment for improved stand productivity. A review of South African research.** *Forest Ecology and Management* 2010, **259**(9):1836-1845.
105. Binkley D, Stape JL: **Sustainable management of Eucalyptus plantations in a changing world.** In: *Eucalyptus in a Changing World-Proceedings of the IUFRO Conference, Aveiro: 2004*; 2004.
106. Graham S, McGinness HM, O'Connell DA: **Effects of management techniques on the establishment of eucalypt seedlings on farmland: a review.** *Agrofor Syst* 2009, **77**(1):59-81.
107. White DA, Crombie DS, Kinal J, Battaglia M, McGrath JF, Mendham DS, Walker SN: **Managing productivity and drought risk in Eucalyptus globulus plantations in south-western Australia.** *Forest Ecology and Management* 2009, **259**(1):33-44.
108. Battaglia M, Bruce J, Brack C, Baker T: **Climate Change and Australia's plantation estate: Analysis of vulnerability and preliminary investigation of adaptation options.** In: *Camberra: Forest & Wood Products Australia*; 2009: 125.
109. van Dijk AIJM, Keenan RJ: **Planted forests and water in perspective.** *Forest Ecology and Management* 2007, **251**(1-2):1-9.
110. Lynch M, Walsh B: **Genetics and analysis of quantitative traits.** Sunderland: Sinauer Associates Incorporated; 1998.
111. Fisher RA: **The Correlation Between Relatives on the Supposition of Mendelian Inheritance.** *Transactions of the Royal Society of Edinburgh* 1918, **52**:399-433.
112. Hamrick JL, Godt M, Sherman-Broyles S: **Factors influencing levels of genetic diversity in woody plant species.** In: *Population Genetics of Forest Trees.* Edited by Adams WT, Strauss S, Copes D, Griffin AR, vol. 42: Springer Netherlands; 1992: 95-124.
113. Caron H, Bandou E, Kremer A: **Multilocus assessment of levels of genetic diversity in tropical trees in Paracou stands.** In: *Ecology and management of a neotropical rainforest : lessons drawn from Paracou, a long-term experimental research site in French Guiana.* Edited by Gourlet Fleury S, Jean marc G, Olivier L, Ecofor GdIPEFP: Elsevier; 2004: 160-170.
114. Snieszko RA, Yanchuk AD, Kliejunas JT, Palmieri KM, Alexander JM, Frankel SJ: **Proceedings of the Fourth International Workshop on the Genetics of Host-Parasite Interactions in Forestry: Disease and Insect Resistance in Forest Trees** Gen. Tech. Rep. PSW-GTR-240. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture.; 2012.
115. Lambeth CC: **Juvenile-mature correlations in Pinaceae and implications for early selection.** *Forest Science* 1980, **26**(4):571-580.
116. Bouvet JM: **Geno-phenotypic regression and juvenile-mature correlations: methodological tools for clonal selection of Eucalyptus hybrids in Congo.** *South African Forestry Journal* 1992(160):13-18.
117. Bouvet JM: **Evolution de la variabilite avec l'age et correlation juvenile-adulte dans des populations d'eucalyptus.** *Ph.D.* Paris: Institut National Agronomique; 1995.
118. Williams DR, Potts BM, Smethurst PJ: **Promotion of flowering in Eucalyptus nitens by paclobutrazol was enhanced by nitrogen fertilizer.** *Canadian Journal of Forest Research* 2003, **33**(1):74-81.
119. Moncur MW, Hasan O: **Floral induction in Eucalyptus nitens.** *Tree Physiology* 1994, **14**(11):1303-1312.



120. Cornelius J: **Heritabilities and additive genetic coefficients of variation in forest trees.** *Canadian Journal of Forest Research* 1994, **24**(2):372-379.
121. Raymond C: **Genetics of Eucalyptus wood properties.** *Ann For Sci* 2002, **59**(5-6):525-531.
122. Bouvet JM, Saya A, Vigneron P: **Trends in additive, dominance and environmental effects with age for growth traits in Eucalyptus hybrid populations.** *Euphytica* 2009, **165**(1):35-54.
123. Osorio LF, White TL, Huber DA: **Age trends of heritabilities and genotype-by-environment interactions for growth traits and wood density from clonal trials of Eucalyptus grandis Hill ex Maiden.** *Silvae genetica* 2001, **50**(1):30-37.
124. Grattapaglia D, Plomion C, Kirst M, Sederoff RR: **Genomics of growth traits in forest trees.** *Current opinion in plant biology* 2009, **12**(2):148-156.
125. Grattapaglia D: **Marker-assisted selection in Eucalyptus.** Rome, Italy: Food and Agriculture Organization of the United Nations (FAO); 2007.
126. Grattapaglia D, Kirst M: **Eucalyptus applied genomics: from gene sequences to breeding tools.** *New Phytol* 2008, **179**(4):911-929.
127. Plomion C, Bousquet J, Kole C: **Genetics, genomics and breeding of conifers:** Science Pub Incorporated; 2011.
128. Kole C: **Genome Mapping and Molecular Breeding in Plants: Forest trees**, vol. 7. Berlin: Springer; 2007.
129. Thavamanikumar S, Southerton SG, Bossinger G, Thumma BR: **Dissection of complex traits in forest trees - opportunities for marker-assisted selection.** *Tree Genetics & Genomes* 2013, **9**(3):627-639.
130. Barton MK: **Twenty years on: The inner workings of the shoot apical meristem, a developmental dynamo.** *Dev Biol* 2010, **341**(1):95-113.
131. Lachaud S, Catesson A-M, Bonnemain J-L: **Structure and functions of the vascular cambium.** *Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie* 1999, **322**(8):633-650.
132. Dye PJ: **Water use efficiency in South African Eucalyptus plantations: A review.** *The Southern African Forestry Journal* 2000, **189**(1):17-26.
133. Laclau J-P, Bouillet J-P, Ranger J: **Dynamics of biomass and nutrient accumulation in a clonal plantation of Eucalyptus in Congo.** *Forest Ecology and Management* 2000, **128**(3):181-196.
134. Laclau J-P, Ranger J, de Moraes Gonçalves JL, Maquère V, Krusche AV, M'Bou AT, Nouvellon Y, Saint-André L, Bouillet J-P, de Cassia Piccolo M *et al*: **Biogeochemical cycles of nutrients in tropical Eucalyptus plantations: Main features shown by intensive monitoring in Congo and Brazil.** *Forest Ecology and Management* 2010, **259**(9):1771-1785.
135. Hubbard RM, Stape J, Ryan MG, Almeida AC, Rojas J: **Effects of irrigation on water use and water use efficiency in two fast growing Eucalyptus plantations.** *Forest Ecology and Management* 2010, **259**(9):1714-1721.
136. Hsiao TC, Acevedo E, Fereres E, Henderson DW: **Water Stress, Growth, and Osmotic Adjustment.** *Philosophical Transactions of the Royal Society of London B, Biological Sciences* 1976, **273**(927):479-500.
137. Downes GM, Drew D, Battaglia M, Schulze D: **Measuring and modelling stem growth and wood formation: An overview.** *Dendrochronologia* 2009, **27**(2):147-157.
138. Stape JL, Binkley D, Ryan MG, Fonseca S, Loos RA, Takahashi EN, Silva CR, Silva SR, Hakamada RE, Ferreira JMD *et al*: **The Brazil Eucalyptus Potential Productivity Project: Influence of water, nutrients and stand uniformity on wood production.** *Forest Ecology and Management* 2010, **259**(9):1684-1694.
139. Villar E: **Plasticité phénotypique et moléculaire de deux clones d'eucalyptus sous contrainte hydrique au champ.** *Ph.D.* Bordeaux: Université de Bordeaux I; 2011.
140. Pienaar LV, Turnbull KJ: **The Chapman-Richards Generalization of Von Bertalanffy's Growth Model for Basal Area Growth and Yield in Even - Aged Stands.** *Forest Science* 1973, **19**(1):2-22.



141. Weiskittel AR, Hann DW, Kershaw Jr JA, Vanclay JK: **Forest growth and yield modeling**: Wiley; 2011.
142. Vanclay JK: **Growth Models for Tropical Forests: A Synthesis of Models and Methods**. *Forest Science* 1995, **41**(1):7-42.
143. Zimmerman D, Núñez-Antón V, Gregoire T, Schabenberger O, Hart J, Kenward M, Molenberghs G, Verbeke G, Pourahmadi M, Vieu P *et al*: **Parametric modelling of growth curve data: An overview**. *Test* 2001, **10**(1):1-73.
144. Mäkelä A, Landsberg J, Ek AR, Burk TE, Ter-Mikaelian M, Ågren GI, Oliver CD, Puttonen P: **Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation**. *Tree Physiology* 2000, **20**(5-6):289-298.
145. Landsberg J: **Physiology in forest models: history and the future**. *FBMIS* 2003, **1**:49-63.
146. Simonneau t, Habib r, Goutouly j-p, Hugué j-g: **Diurnal Changes in Stem Diameter Depend Upon Variations in Water Content: Direct Evidence in Peach Trees**. *Journal of Experimental Botany* 1993, **44**(3):615-621.
147. Herzog K, Häsler R, Thum R: **Diurnal changes in the radius of a subalpine Norway spruce stem: their relation to the sap flow and their use to estimate transpiration**. *Trees* 1995, **10**(2):94-101.
148. Zweifel R, Item H, Häsler R: **Link between diurnal stem radius changes and tree water relations**. *Tree Physiology* 2001, **21**(12-13):869-877.
149. Cochard H, Forestier S, Améglio T: **A new validation of the Scholander pressure chamber technique based on stem diameter variations**. *Journal of Experimental Botany* 2001, **52**(359):1361-1365.
150. Garnier E, Berger A: **Effect of Water Stress on Stem Diameter Changes of Peach Trees Growing in the Field**. *J Appl Ecol* 1986, **23**(1):193-209.
151. Sevanto S, Vesala T, Perämäki M, Nikinmaa E: **Time lags for xylem and stem diameter variations in a Scots pine tree**. *Plant, Cell & Environment* 2002, **25**(8):1071-1077.
152. Zweifel R, Zimmermann L, Newbery DM: **Modeling tree water deficit from microclimate: an approach to quantifying drought stress**. *Tree Physiology* 2005, **25**(2):147-156.
153. Craig H: **Isotopic standards for carbon and oxygen and correction factors for mass-spectrometric analysis of carbon dioxide**. *Geochimica et Cosmochimica Acta* 1957, **12**(1-2):133-149.
154. Farquhar G, Richards R: **Isotopic Composition of Plant Carbon Correlates With Water-Use Efficiency of Wheat Genotypes**. *Functional Plant Biology* 1984, **11**(6):539-552.
155. O'Leary MH: **Carbon isotope fractionation in plants**. *Phytochemistry* 1981, **20**(4):553-567.
156. Farquhar G, O'Leary M, Berry J: **On the Relationship Between Carbon Isotope Discrimination and the Intercellular Carbon Dioxide Concentration in Leaves**. *Functional Plant Biology* 1982, **9**(2):121-137.
157. Farquhar G, Ehleringer J, Hubick K: **Carbon Isotope Discrimination and Photosynthesis**. *Annual Review of Plant Physiology and Plant Molecular Biology* 1989, **40**(1):503-537.
158. Farquhar G, Hubick K, Condon A, Richards R: **Carbon Isotope Fractionation and Plant Water-Use Efficiency**. In: *Stable Isotopes in Ecological Research*. Edited by Rundel PW, Ehleringer JR, Nagy KA, vol. 68: Springer New York; 1989: 21-40.
159. Cernusak LA, Tcherkez G, Keitel C, Cornwell WK, Santiago LS, Knohl A, Barbour MM, Williams DG, Reich PB, Ellsworth DS *et al*: **Why are non-photosynthetic tissues generally <sup>13</sup>C enriched compared with leaves in C3 plants? Review and synthesis of current hypotheses**. *Functional Plant Biology* 2009, **36**(3):199-213.
160. Le Roux X, Bariac T, Sinoquet H, Genty B, Piel C, Mariotti A, Girardin C, Richard P: **Spatial distribution of leaf water-use efficiency and carbon isotope discrimination within an isolated tree crown**. *Plant, Cell & Environment* 2001, **24**(10):1021-1032.
161. Wullschlegel SD, Meinzer FC, Vertessy RA: **A review of whole-plant water use studies in tree**. *Tree Physiology* 1998, **18**(8-9):499-512.





162. Ehleringer JR, Hall AE, Farquhar GD: **Stable isotopes and plant carbon - water relations**. San Diego: Academic Press; 1993.
163. Hall AE, Richards RA, Condon AG, Wright GC, Farquhar GD: **Carbon Isotope Discrimination and Plant Breeding**. In: *Plant Breeding Reviews*. John Wiley & Sons, Inc.; 2010: 81-113.
164. Warren C, McGrath J, Adams M: **Water availability and carbon isotope discrimination in conifers**. *Oecologia* 2001, **127**(4):476-486.
165. Macfarlane C, Adams MA:  **$\delta^{13}\text{C}$  of wood in growth-rings indicates cambial activity of drought-stressed trees of *Eucalyptus globulus***. *Functional Ecology* 1998, **12**(4):655-664.
166. Nguyen-Queyrens A, Ferhi A, Loustau D, Guehl J-M: **Within-ring  $\delta^{13}\text{C}$  spatial variability and interannual variations in wood cellulose of two contrasting provenances of *Pinus pinaster***. *Canadian Journal of Forest Research* 1998, **28**(5):766-773.
167. Brendel O, Pot D, Plomion C, Rozenberg P, Guehl JM: **Genetic parameters and QTL analysis of  $\delta^{13}\text{C}$  and ring width in maritime pine**. *Plant, Cell & Environment* 2002, **25**(8):945-953.
168. Prasolova NV, Xu ZH, Lundkvist K, Farquhar GD, Dieters MJ, Walker S, Saffigna PG: **Genetic variation in foliar carbon isotope composition in relation to tree growth and foliar nitrogen concentration in clones of the F1 hybrid between slash pine and Caribbean pine**. *Forest Ecology and Management* 2003, **172**(2-3):145-160.
169. Monclus R, Villar M, Barbaroux C, Bastien C, Fichot R, Delmotte FM, Delay D, Petit J-M, Br  chet C, Dreyer E *et al*: **Productivity, water-use efficiency and tolerance to moderate water deficit correlate in 33 poplar genotypes from a *Populus deltoides*  $\times$  *Populus trichocarpa* F1 progeny**. *Tree Physiology* 2009, **29**(11):1329-1339.
170. Xu Z, Prasolova N, Lundkvist K, Beadle C, Leaman T: **Genetic variation in branchlet carbon and nitrogen isotope composition and nutrient concentration of 11-year-old hoop pine families in relation to tree growth in subtropical Australia**. *Forest Ecology and Management* 2003, **186**(1-3):359-371.
171. Schulze E-D, Turner NC, Nicolle D, Schumacher J: **Leaf and wood carbon isotope ratios, specific leaf areas and wood growth of *Eucalyptus* species across a rainfall gradient in Australia**. *Tree Physiology* 2006, **26**(4):479-492.
172. Monclus R, Dreyer E, Villar M, Delmotte FM, Delay D, Petit J-M, Barbaroux C, Le Thiec D, Br  chet C, Brignolas F: **Impact of drought on productivity and water use efficiency in 29 genotypes of *Populus deltoides* $\times$ *Populus nigra***. *New Phytol* 2006, **169**(4):765-777.
173. Chamaillard S, Fichot R, Vincent-Barbaroux C, Bastien C, Depierreux C, Dreyer E, Villar M, Brignolas F: **Variations in bulk leaf carbon isotope discrimination, growth and related leaf traits among three *Populus nigra* L. populations**. *Tree Physiology* 2011, **31**(10):1076-1087.
174. Pita P, Soria F, Ca  as I, Toval G, Pardos JA: **Carbon isotope discrimination and its relationship to drought resistance under field conditions in genotypes of *Eucalyptus globulus* Labill**. *Forest Ecology and Management* 2001, **141**(3):211-221.
175. Macfarlane C, Adams MA, White DA: **Productivity, carbon isotope discrimination and leaf traits of trees of *Eucalyptus globulus* Labill. in relation to water availability**. *Plant, Cell & Environment* 2004, **27**(12):1515-1524.
176. Rowell DM, Ades PK, Tausz M, Arndt SK, Adams MA: **Lack of genetic variation in tree ring  $\delta^{13}\text{C}$  suggests a uniform, stomatally-driven response to drought stress across *Pinus radiata* genotypes**. *Tree Physiology* 2009, **29**(2):191-198.
177. Leavitt SW, Long A: **Stable-Carbon Isotope Variability in Tree Foliage and Wood**. *Ecology* 1986, **67**(4):1002-1010.
178. Lambeth C, Endo M, Wright J: **Genetic Analysis of 16 Clonal Trials of *Eucalyptus grandis* and Comparisons with Seedling Checks**. *Forest Science* 1994, **40**(3):397-411.
179. Hannrup B, Wilhelmsson L, Danell O: **Time trends for genetic parameters of wood density and growth traits in *Pinus sylvestris* L.** *Silvae genetica* 1998, **47**(4):214-219.



180. Franklin EC: **Model relating levels of genetic variance to stand development of 4 north-american conifers.** *Silvae genetica* 1979, **28**(5-6):207-212.
181. Bouvet JM, Vigneron P: **Age trends in variances and heritabilities in Eucalyptus factorial mating designs.** *Silvae genetica* 1995, **44**(4):206-216.
182. Stackpole D, Vaillancourt R, de Aguiar M, Potts B: **Age trends in genetic parameters for growth and wood density in Eucalyptus globulus.** *Tree Genetics & Genomes* 2010, **6**(2):179-193.
183. Jansson G, Li B, Hannrup B: **Time Trends in Genetic Parameters for Height and Optimal Age for Parental Selection in Scots Pine.** *Forest Science* 2003, **49**(5):696-705.
184. Danjon F: **Heritabilities and genetic correlations for estimated growth curve parameters in maritime pine.** *TAG Theoretical and Applied Genetics* 1994, **89**(7):911-921.
185. Gwaze DG, Bridgwater FB, Williams CW: **Genetic analysis of growth curves for a woody perennial species, Pinus taeda.** *TAG Theoretical and Applied Genetics* 2002, **105**(4):526-531.
186. Neale DB, Williams CG: **Restriction fragment length polymorphism mapping in conifers and applications to forest genetics and tree improvement.** *Canadian Journal of Forest Research* 1991, **21**(5):545-554.
187. Grattapaglia D, Sederoff R: **Genetic Linkage Maps of Eucalyptus grandis and Eucalyptus urophylla Using a Pseudo-Testcross: Mapping Strategy and RAPD Markers.** *Genetics* 1994, **137**(4):1121-1137.
188. Myburg AA, Potts BM, Marques CM, Kirst M, Gion JM, Grattapaglia D, Grima-Pettenatti J: **Eucalypts.** Heidelberg, Germany: Springer-Verlag GmbH; 2007.
189. Kaya Z, Sewell MM, Neale DB: **Identification of quantitative trait loci influencing annual height- and diameter-increment growth in loblolly pine (Pinus taeda L.).** *Theor Appl Genet* 1999, **98**(3-4):586-592.
190. Emebiri LC, Devey ME, Matheson AC, Slee MU: **Age-related changes in the expression of QTLs for growth in radiata pine seedlings.** *TAG Theoretical and Applied Genetics* 1998, **97**(7):1053-1061.
191. Freeman JS, Whittock SP, Potts BM, Vaillancourt RE: **QTL influencing growth and wood properties in Eucalyptus globulus.** *Tree Genetics & Genomes* 2009, **5**(4):713-722.
192. Grattapaglia D, Bertolucci F, Penchel R, Sederoff R: **Genetic mapping of quantitative trait loci controlling growth and wood quality traits in Eucalyptus grandis using a maternal half-sib family and RAPD markers.** *Genetics* 1996, **144**:1205 - 1214.
193. Verhaegen D, Plomion C, Gion JM, Poitel M, Costa P, Kremer A: **Quantitative trait dissection analysis in Eucalyptus using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages.** *TAG Theoretical and Applied Genetics* 1997, **95**(4):597-608.
194. Beavis WD: **QTL analyses: power, precision, and accuracy.** *Molecular dissection of complex traits* 1998, **1998**:145-162.
195. Rönnerberg-Wästljung AC, Glynn C, Weih M: **QTL analyses of drought tolerance and growth for a Salix dasyclados × Salix viminalis hybrid in contrasting water regimes.** *TAG Theoretical and Applied Genetics* 2005, **110**(3):537-549.
196. Weih M, Rönnerberg-Wästljung AC, Glynn C: **Genetic basis of phenotypic correlations among growth traits in hybrid willow (Salix dasyclados x S-viminalis) grown under two water regimes.** *New Phytol* 2006, **170**(3):467-477.
197. Wu R, Bradshaw HD, Stettler RF: **Developmental quantitative genetics of growth in Populus.** *Theor Appl Genet* 1998, **97**(7):1110-1119.
198. Rae A, Pinel M, Bastien C, Sabatti M, Street N, Tucker J, Dixon C, Marron N, Dillen S, Taylor G: **QTL for yield in bioenergy Populus: identifying G×E interactions from growth at three contrasting sites.** *Tree Genetics & Genomes* 2008, **4**(1):97-112.
199. Rae AM, Street NR, Robinson KM, Harris N, Taylor G: **Five QTL hotspots for yield in short rotation coppice bioenergy poplar: The Poplar Biomass Loci.** *BMC Plant Biology* 2009, **9**.



200. Pelgas B, Bousquet J, Meirmans P, Ritland K, Isabel N: **QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments.** *BMC Genomics* 2011, **12**(1):145.
201. Freeman JS, Potts BM, Downes GM, Pilbeam D, Thavamanikumar S, Vaillancourt RE: **Stability of quantitative trait loci for growth and wood properties across multiple pedigrees and environments in *Eucalyptus globulus*.** *New Phytol* 2013, **198**(4):1121-1134.
202. Bundock PC, Potts BM, Vaillancourt RE: **Detection and stability of quantitative trait loci (QTL) in *Eucalyptus globulus*.** *Tree Genetics & Genomes* 2008, **4**(1):85-95.
203. Lerceteau E, Szmidt AE, Andersson B: **Detection of quantitative trait loci in *Pinus sylvestris* L. across years.** *Euphytica* 2001, **121**(2):117-122.
204. Thumma BR, Baltunis BS, Bell JC, Emebiri LC, Moran GF, Southerton SG: **Quantitative trait locus (QTL) analysis of growth and vegetative propagation traits in *Eucalyptus nitens* full-sib families.** *Tree Genetics & Genomes* 2010, **6**(6):877-889.
205. Sansaloni CP, Petroli CD, Carling J, Hudson CJ, Steane DA, Myburg AA, Grattapaglia D, Vaillancourt RE, Kilian A: **A high-density Diversity Arrays Technology (DArT) microarray for genome-wide genotyping in *Eucalyptus*.** *Plant Methods* 2010, **6**.
206. Neves L, MC Mamani E, Alfenas A, Kirst M, Grattapaglia D: **A high-density transcript linkage map with 1,845 expressed genes positioned by microarray-based Single Feature Polymorphisms (SFP) in *Eucalyptus*.** *BMC Genomics* 2011, **12**(1):189.
207. Grattapaglia D, Silva-Junior O, Kirst M, de Lima B, Faria D, Pappas G: **High-throughput SNP genotyping in the highly heterozygous genome of *Eucalyptus*: assay success, polymorphism and transferability across species.** *BMC Plant Biology* 2011, **11**(1):65.
208. Pavy N, Gagnon F, Rigault P, Blais S, Deschenes A, Boyle B, Pelgas B, Deslauriers M, Clement S, Lavigne P *et al*: **Development of high-density SNP genotyping arrays for white spruce (*Picea glauca*) and transferability to subtropical and nordic congeners.** *Molecular Ecology Resources* 2013, **13**(2):324-336.
209. Hudson CJ, Kullam ARK, Freeman JS, Faria DA, Grattapaglia D, Kilian A, Myburg AA, Potts BM, Vaillancourt RE: **High synteny and colinearity among *Eucalyptus* genomes revealed by high-density comparative genetic mapping.** *Tree Genetics & Genomes* 2012, **8**(2):339-352.
210. Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A *et al*: **The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray).** *Science* 2006, **313**(5793):1596-1604.
211. Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin Y-C, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A *et al*: **The Norway spruce genome sequence and conifer genome evolution.** *Nature* 2013, **497**(7451):579-584.
212. van Eeuwijk FA, Malosetti M, Yin X, Struik PC, Stam P: **Statistical models for genotype by environment data: from conventional ANOVA models to eco-physiological QTL models.** *Australian Journal of Agricultural Research* 2005, **56**(9):883-894.
213. van Eeuwijk FA, Bink MCAM, Chenu K, Chapman SC: **Detection and use of QTL for complex traits in multiple environments.** *Current opinion in plant biology* 2010, **13**(2):193-205.
214. Wang Y, Wang N, Wang Z, Han H, Wu R: **Genetic Dissection of Complex Traits: From Functional Mapping to Systems Mapping.** *Journal of Biometrics & Biostatistics* 2012.
215. Wu R, Lin M: **Functional mapping : how to map and study the genetic architecture of dynamic complex traits.** *Nat Rev Genet* 2006, **7**(3):229-237.
216. Emebiri LC: **Detection and genetic mapping of quantitative trait loci influencing stem growth efficiency in radiata pine.** *Ph.D.*: Australian National University; 1997.
217. Jermstad KD, Bassoni DL, Jech KS, Ritchie GA, Wheeler NC, Neale DB: **Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas fir. III. Quantitative trait loci-by-environment interactions.** *Genetics* 2003, **165**(3):1489-1506.



218. Jansen RC, Stam P: **High resolution of quantitative traits into multiple loci via interval mapping.** *Genetics* 1994, **136**(4):1447-1455.
219. Korol AB, Ronin YI, Nevo E: **Approximate Analysis of QTL-Environment Interaction with No Limits on the Number of Environments.** *Genetics* 1998, **148**(4):2015-2028.
220. Sari-Gorla M, Calinski T, Kaczmarek Z, Krajewski P: **Detection of QTL [times] environment interaction in maize by a least squares interval mapping method.** *Heredity* 1997, **78**(2):146-157.
221. Wang DL, Zhu J, Li ZKL, Paterson AH: **Mapping QTLs with epistatic effects and QTL×environment interactions by mixed linear model approaches.** *Theor Appl Genet* 1999, **99**(7-8):1255-1264.
222. Van Eeuwijk F, Crossa J, Vargas M, Ribaut J: **Analysing QTL–Environment Interaction by Factorial Regression, with an Application to the CIMMYT Drought and Low-nitrogen Stress Programme in Maize.** In: *Quantitative Genetics, Genomics and Plant Breeding*. Edited by Kang MS. Wallingford: CABI Publishing; 2002.
223. Verbyla AP, Eckermann PJ, Thompson R, Cullis BR: **The analysis of quantitative trait loci in multi-environment trials using a multiplicative mixed model.** *Australian Journal of Agricultural Research* 2003, **54**(12):1395-1408.
224. Piepho H-P, Pillen K: **Mixed modelling for QTL × environment interaction analysis.** *Euphytica* 2004, **137**(1):147-153.
225. Malosetti M, Voltas J, Romagosa I, Ullrich SE, van Eeuwijk FA: **Mixed models including environmental covariables for studying QTL by environment interaction.** *Euphytica* 2004, **137**(1):139-145.
226. Yandell BS, Mehta T, Banerjee S, Shriner D, Venkataraman R, Moon JY, Neely WW, Wu H, von Smith R, Yi N: **R/qtlbim: QTL with Bayesian Interval Mapping in experimental crosses.** *Bioinformatics* 2007, **23**(5):641-643.
227. Jansen RC, Ooijen JW, Stam P, Lister C, Dean C: **Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci.** *Theor Appl Genet* 1995, **91**(1):33-37.
228. Boer MP, Wright D, Feng L, Podlich DW, Luo L, Cooper M, van Eeuwijk FA: **A Mixed-Model Quantitative Trait Loci (QTL) Analysis for Multiple-Environment Trial Data Using Environmental Covariables for QTL-by-Environment Interactions, With an Example in Maize.** *Genetics* 2007, **177**(3):1801-1813.
229. Malosetti M, Ribaut J-M, van Eeuwijk FA: **The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis.** *Frontiers in Physiology* 2013, **4**.
230. Brendel O, Thiec D, Scotti-Saintagne C, Bodénès C, Kremer A, Guehl J-M: **Quantitative trait loci controlling water use efficiency and related traits in Quercus robur L.** *Tree Genetics & Genomes* 2008, **4**(2):263-278.
231. Walsh B, Lynch M: **Evolution and selection of quantitative traits.** *Foundations Sunderland, MA: Sinauer* 2009.
232. Gauch Jr HG: **Model selection and validation for yield trials with interaction.** *Biometrics* 1988, **44**(3):705-715.
233. Finlay K, Wilkinson G: **The analysis of adaptation in a plant-breeding programme.** *Australian Journal of Agricultural Research* 1963, **14**(6):742-754.
234. Reymond M, Muller B, Tardieu F: **Dealing with the genotype×environment interaction via a modelling approach: a comparison of QTLs of maize leaf length or width with QTLs of model parameters.** *Journal of Experimental Botany* 2004, **55**(407):2461-2472.
235. Li Q, Huang Z, Xu M, Wang C, Gai J, Huang Y, Pang X, Wu R: **Functional mapping of genotype-environment interactions for soybean growth by a semiparametric approach.** *Plant Methods* 2010, **6**(1):13.
236. Stratton DA: **Reaction norm functions and QTL-environment interactions for flowering time in Arabidopsis thaliana.** *Heredity* 1998, **81**(2):144-155.



237. Reymond M, Muller B, Leonardi A, Charcosset A, Tardieu F: **Combining Quantitative Trait Loci Analysis and an Ecophysiological Model to Analyze the Genetic Variability of the Responses of Maize Leaf Growth to Temperature and Water Deficit.** *Plant Physiol* 2003, **131**(2):664-675.
238. Yin X, Struik PC, van Eeuwijk FA, Stam P, Tang J: **QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley.** *Journal of Experimental Botany* 2005, **56**(413):967-976.
239. Ma C-X, Casella G, Wu R: **Functional Mapping of Quantitative Trait Loci Underlying the Character Process: A Theoretical Framework.** *Genetics* 2002, **161**(4):1751-1762.
240. Li Y, Wu R: **Functional mapping of growth and development.** *Biological Reviews* 2010, **85**(2):207-216.
241. Ma CX, Wu R, Casella G: **FunMap: functional mapping of complex traits.** *Bioinformatics* 2004, **20**(11):1808-1811.
242. Wu R, Ma C-X, Chang M, Littell RC, Wu SS, Yin T, Huang M, Wang M, Casella G: **A logistic mixture model for characterizing genetic determinants causing differentiation in growth trajectories.** *Genetics Research* 2002, **79**(03).
243. Wu R, Ma C-X, Lin M, Casella G: **A General Framework for Analyzing the Genetic Architecture of Developmental Characteristics.** *Genetics* 2004, **166**(3):1541-1551.
244. Wu R, Ma C-X, Lin M, Wang Z, Casella G: **Functional Mapping of Quantitative Trait Loci Underlying Growth Trajectories Using a Transform-Both-Sides Logistic Model.** *Biometrics* 2004, **60**(3):729-738.
245. Wu RL, Ma CX, Zhao W, Casella G: **Functional mapping for quantitative trait loci governing growth rates: a parametric model.** *Physiol Genomics* 2003, **14**(3):241-249.
246. Malosetti M, Visser R, Celis-Gamboa C, van Eeuwijk F: **QTL methodology for response curves on the basis of non-linear mixed models, with an illustration to senescence in potato.** *TAG Theoretical and Applied Genetics* 2006, **113**(2):288-300.
247. Zhang B, Tong CF, Yin TM, Zhang XY, Zhuge QQ, Huang MR, Wang MX, Wu RL: **Detection of quantitative trait loci influencing growth trajectories of adventitious roots in Populus using functional mapping.** *Tree Genetics & Genomes* 2009, **5**(3):539-552.
248. Wu S, Yang J, Huang Y, Li Y, Yin T, Wulschleger SD, Tuskan GA, Wu R: **An Improved Approach for Mapping Quantitative Trait Loci in a Pseudo-Testcross: Revisiting a Poplar Mapping Study.** *Bioinformatics and Biology Insights* 2010, **4**(1832-BBI-An-Improved-Approach-for-Mapping-Quantitative-Trait-Loci-in-a-Pseudo-T.pdf):1-8.





## **2 Matériel et méthodes**



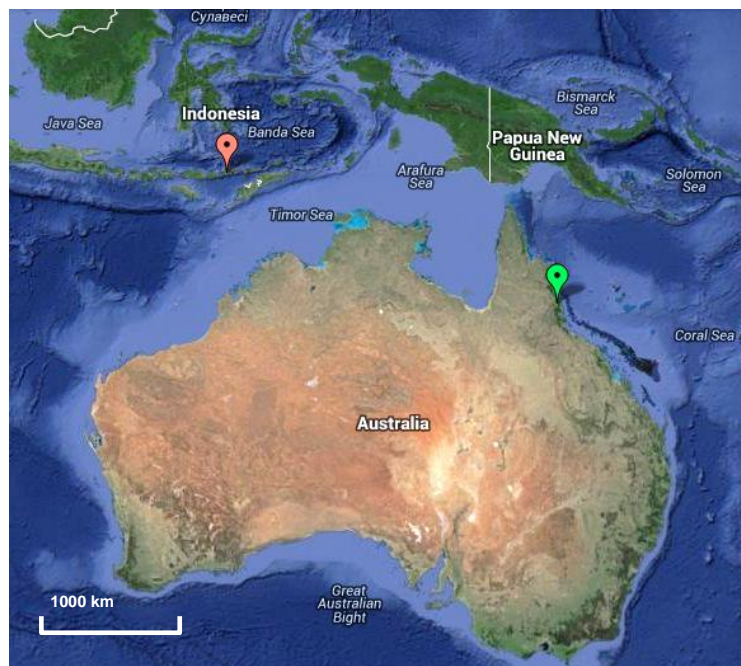


Figure 2.1 : Localisation des provenances des deux géniteurs du croisement étudié : *E. urophylla* en rouge et *E. grandis* en vert.





Cette partie est consacrée à la description des quatre essais (P93, P97, P98 et P10) étudiés au cours de cette thèse. Elle présente les différentes données utilisées ainsi que l'ensemble des méthodes d'analyse mises en œuvre. Les méthodes plus spécifiques sont abordées dans les parties "matériels et méthodes" des articles. Le dispositif de suivi en continu des micro-variations du rayon (P10), dont la mise en place a demandé un investissement important en première année de thèse, est présenté de manière plus détaillée dans ce chapitre.

## 2.1 Matériel végétal et essais

Le matériel végétal étudié au cours de cette thèse est un croisement interspécifique entre deux génotypes d'*Eucalyptus* : un génotype d'*E. urophylla* (14.144) utilisé comme parent femelle et un génotype d'*E. grandis* (9.21) utilisé comme parent mâle. Cette combinaison hybride est l'une des plus performantes dans les conditions congolaises. Elle est intégrée comme famille élite dans le programme d'amélioration génétique mené par le CRDPI (Centre de Recherche sur la Durabilité et la Productivité des Plantations Industrielles) en République du Congo. Ce croisement (14.144 x 9.21) a donc été choisi pour cette thèse, car il constitue le croisement de référence au CIRAD pour les travaux de cartographie génétique et d'analyse de l'architecture génétique des caractères de croissance et de qualité du bois [1-4].

Ces deux espèces sont très proches phylogénétiquement [5, 6], elles appartiennent au même sous-genre botanique (*Symphyomyrtus*) et à la même section (*Latoangulatae*). Leur hybridation est rendue possible par cette proximité génétique [7]. En revanche, leurs aires de répartition naturelles sont géographiquement séparées.

*E. urophylla* occupe l'archipel des petites îles de la Sonde partagé entre l'Indonésie et le Timor oriental. On la rencontre à des altitudes très variables (de 300 à 3000 m) et sous des régimes de précipitations contrastés (de 3 à 8 mois de saison sèche) [8]. Le génotype d'*E. urophylla* 14.144 provient de l'île de Florès en Indonésie à proximité du volcan Lewotobi (122°46'E 8°31'S, Figure 2.1). *E. grandis* se trouve exclusivement sur la côte Est de l'Australie où les précipitations annuelles moyennes varient entre 1100 et 3500 mm, avec des températures annuelles moyennes de 12 à 24°C [9]. Le génotype *E. grandis* 9.21 provient de la localité d'Atherton (145°40'E 17°17'S, Figure 2.1).



Tableau 2.1: Caractéristiques des quatre essais étudiés. Ht : hauteur, C : circonférence, D : diamètre

Essai		P93	P97	P98	P10
Date de plantation		avr.-1993	avr.-1997	nov.-1998	nov.-2010
Densité (arbres/ha)		667	667	802	802
Données phénotypiques	Nombre d'individus	201	600	600 (60 x 10 clones)	2500 dont 240
	Mesure de croissance	Ht / C	Ht / C	Ht / C	Ht / D micro-variations
	Nombre de mesures	5	17	10	7 288 par jour
	Age final	59 mois	62 mois	60 mois	- -
	Mesure de $\delta^{13}\text{C}$	1 sur bois	1 sur bois	-	- 2 sur feuilles
Données génotypiques	Nombre d'individus marqueurs	201	190	60	1040
		769 +5175	278	769 +5175	5175
	Type de marqueurs	RAPD, SSR, EST, SNP	RAPD	RAPD - SSR - EST - SNP	SNP
Données climatiques	Lieu	Station de Pointe Noire			Sur l'essai
	Pas de temps	Journalier			5 min
	Variables mesurées	T°C, pluviométrie, humidité relative de l'air (HR)			T°C, pluviométrie, HR, humidité du sol, rayonnement incident

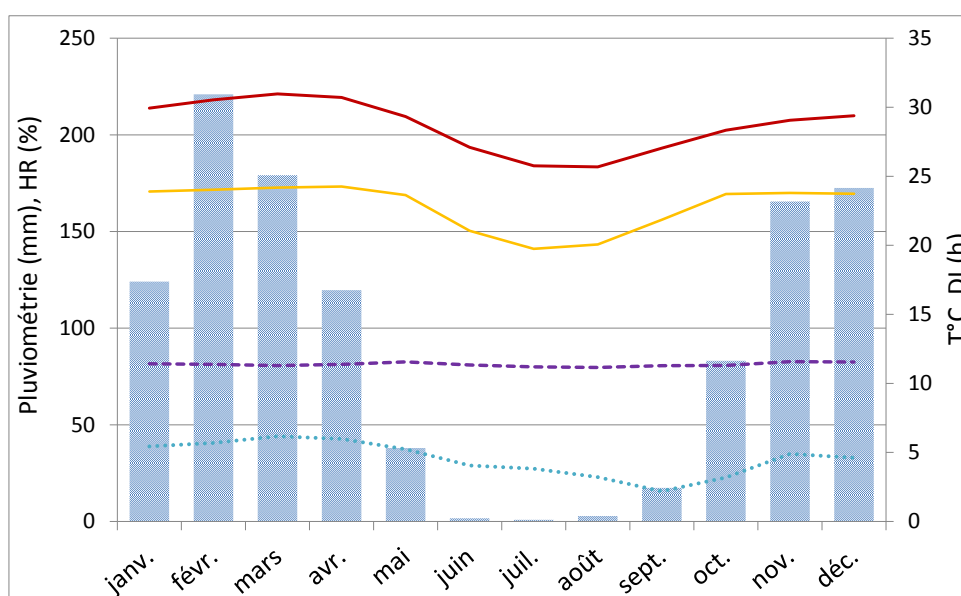


Figure 2.2: Normales mensuelles pour la station météo de Pointe-Noire en République du Congo (calculées sur la période : 1988 – 2006). Précipitations (barres bleues), HR : humidité relative (tirets violets), T°C : température (maximum en rouge, minimum en jaune), DI : durée d'insolation moyenne par jour (pointillés bleus).



Ces deux espèces ont montré des comportements contrastés en République du Congo [10], que l'on cherche à combiner au sein d'hybrides interspécifiques. En effet, l'espèce *E. grandis* qui présente de fortes potentialités de croissance est peu adaptée aux conditions congolaises, contrairement à l'espèce *E. urophylla*. Cette adaptation d'*E. urophylla* se manifeste, entre autres, par une floraison importante (nombre de fleurs et durée de la floraison) et de meilleure qualité pour réaliser des croisements contrôlés (fleurs plus grosses et moins fragiles qu'*E. grandis*), ce qui a conduit à son utilisation comme parent femelle.

Des descendants hybrides (F'1) issus du croisement *E. urophylla* (14.144) x *E. grandis* (9.21) ont été plantés dans quatre dispositifs différents (Tableau 2.1) : trois sont antérieurs à cette thèse (dispositifs P93, P97 et P98 plantés respectivement en 1993, 1997 et 1998) et le quatrième (P10) a été installé au démarrage de cette thèse en novembre 2010. P93, P97 et P10 correspondent à des dispositifs QTL plantés avec des génotypes différents (familles de plein-frères sans répétitions) alors que P98 est un test clonal de 60 génotypes issus de P93, avec 10 répétitions clonales.

L'installation et le suivi des essais ainsi que les mesures phénotypiques ont été réalisés en partenariat avec le CRDPI localisé à Pointe-Noire (République du Congo).

## 2.2 Site d'étude et conditions pédoclimatiques

Le site d'étude se situe en République du Congo dans la région de Pointe-Noire. Le climat de cette région est classifié comme "savane équatoriale avec hiver sec" [11]. Il présente deux saisons contrastées en termes de disponibilité en eau : la saison des pluies d'octobre à mai, et la saison sèche de juin à septembre. Les précipitations moyennes sont de 1200 mm par an, l'humidité relative de l'air moyenne est de 85% et la température moyenne est de 25°C. Ces valeurs moyennes cachent des variations interannuelles parfois fortes pour les précipitations et la durée relative des deux saisons. Contrairement aux précipitations, la température et l'humidité ne présentent que de faibles variations saisonnières : 2% pour l'humidité de l'air et 3°C pour la température (Figure 2.2). Ces conditions climatiques sont donc favorables pour observer l'effet des variations de la disponibilité en eau sur la croissance au champ des eucalyptus.

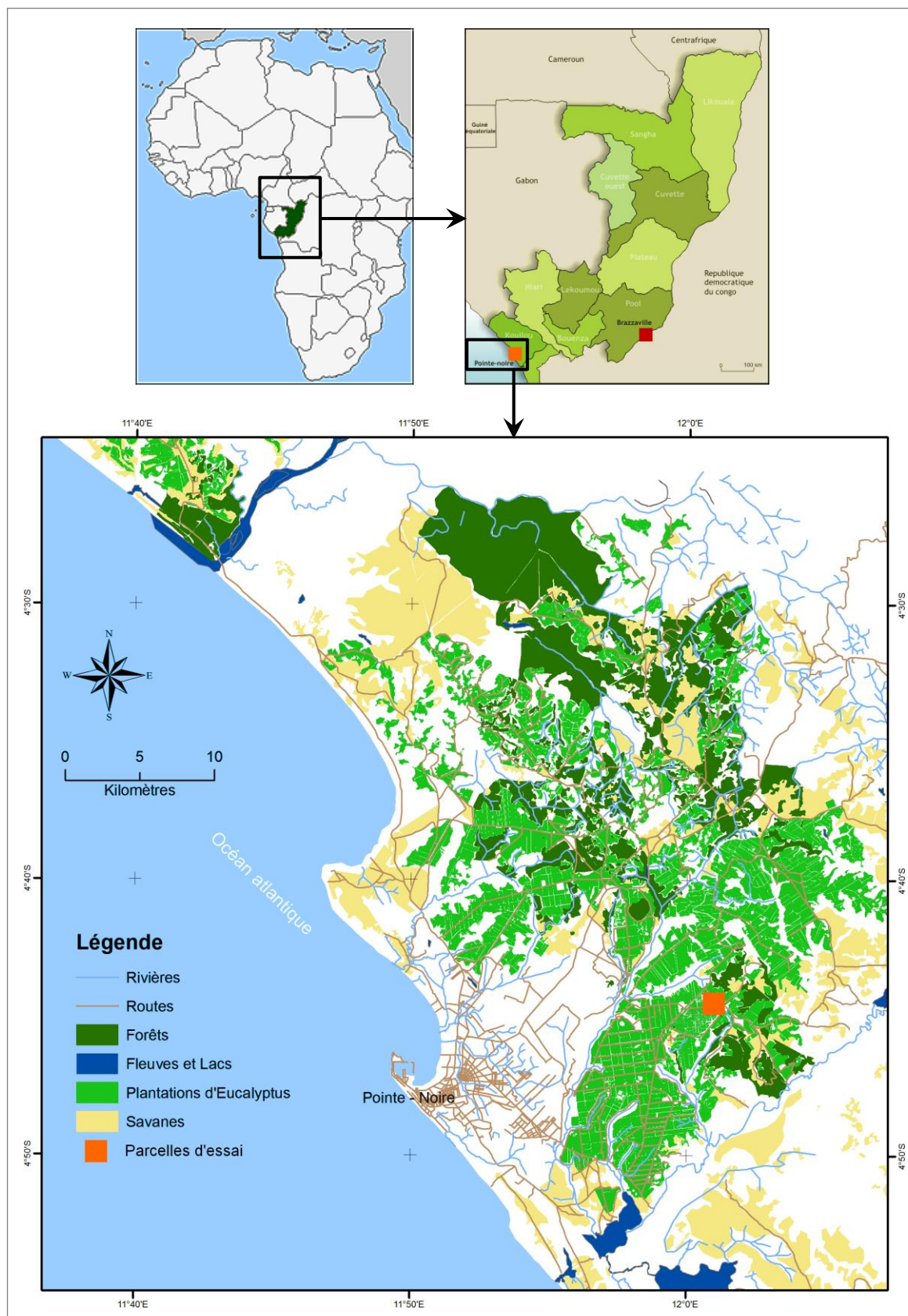


Figure 2.3 : Localisation des parcelles au sein du massif forestier d'*Eucalyptus* dans la région côtière de Pointe-Noire (République du Congo). La carte de la région de Pointe-Noire a été réalisée par A.S. Dzomambou au CRDPI.



Les quatre essais ont été plantés sur le plateau de Kissoko, à l'est de Pointe-Noire (4.45° S, 12° E, Figure 2.3). Ce plateau, formé par l'érosion de la chaîne de montagnes du Mayombe, située à 100 km à l'Est, est caractérisé par des sols très profonds, principalement composés de sables ( $\approx 90\%$ ). Ces sols sont considérés comme pauvres, du fait de leur faible rétention en eau et d'une faible capacité d'échange cationique [12]. Ils sont classés en tant qu'Arenosols ferraliques [13] et sont naturellement recouverts par de la savane herbacée.

## 2.3 Installation, suivi et traitement du dispositif P10

Le dispositif P10 a été installé dans le cadre du projet FEDER Abiogen (N° 32973), qui porte sur l'analyse de l'architecture génétique de la dynamique de croissance de l'*Eucalyptus* (Pointe-Noire) et du Pin maritime (Bordeaux). Chaque dispositif permet le suivi en continu des micro-variations du rayon sur 240 arbres grâce à des micro-dendromètres, ainsi qu'un suivi de variables environnementales comme la température ou l'humidité de l'air et du sol. Si ce type de suivi dynamique n'est pas nouveau chez les arbres [14-17], sa mise en place à grande échelle (240 individus suivis par dispositif) n'avait jamais été réalisée avant cette thèse. Leur installation a donc nécessité de lever des verrous techniques qui sont détaillés dans cette partie pour le cas de l'*Eucalyptus*.

### 2.3.1 Conception du dispositif

#### 2.3.1.1 Dispositif initial

Le plan théorique du dispositif a été pensé avec quatre placeaux instrumentés, autonomes au niveau de l'acquisition des données et de l'alimentation électrique. Il était prévu de mesurer les micro-variations du rayon de 64 arbres par placeau ainsi que la température et l'humidité de l'air et du sol. Un placeau est équipé d'une centrale de mesure, d'un capteur de température et d'humidité de l'air et de quatre sondes de mesure de l'humidité du sol. Un placeau comporte également deux multiplexeurs afin de démultiplier la capacité en micro-dendromètre, et donc le nombre d'arbres pouvant être suivis. Chaque placeau est alimenté par une batterie reliée à un panneau solaire pour assurer l'autonomie électrique. Cette configuration théorique a été choisie afin d'optimiser le nombre de centrales de mesure, tout en maintenant une faible distance entre l'arbre et la centrale. L'autonomie d'un placeau permet aussi de minimiser la perte de données en cas d'incident.

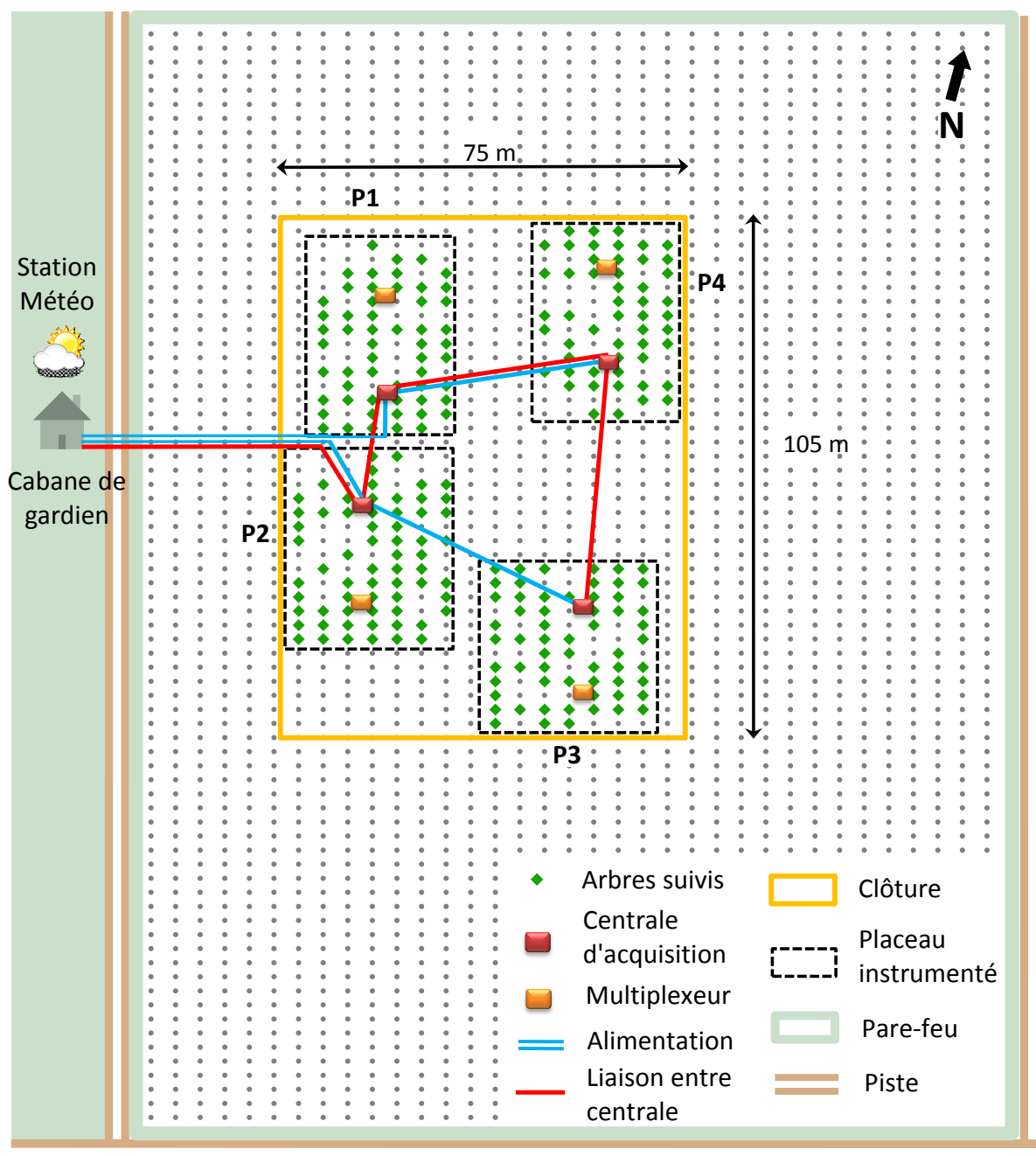


Figure 2.4: Plan de l'essai P10 avec les infrastructures et la localisation des arbres suivis sur les quatre placeau (P1 à P4)





### 2.3.1.2 Adaptation au cas de l'eucalyptus

Cette configuration théorique a ensuite été adaptée aux contextes "pin maritime" puis "eucalyptus". En effet, l'installation du dispositif "pin maritime" de novembre 2010 à début mars 2011, a permis à l'ensemble de l'équipe (Didier Bert, Régis Burlett, Jean-Marc Gion, André Mabiala et moi-même) d'appréhender les difficultés techniques. Cela a permis d'installer de manière efficace, sur une courte durée (1 mois), le dispositif "eucalyptus" à Pointe-Noire. En parallèle, une mission de dix jours au Congo en février 2011, a permis d'apporter du matériel sur place et de préciser les plans du dispositif "eucalyptus".

Afin d'assurer la sécurité du dispositif, il a été décidé d'installer une cabane de gardiennage à proximité immédiate de l'essai, ainsi qu'une clôture autour des placeaux suivis. Il a donc fallu adapter l'alimentation électrique du dispositif pour permettre un positionnement excentré (cabane de gardien) des batteries et des panneaux solaires (Figure 2.4). Il a également été décidé de mettre en place un système de communication entre les centrales des quatre placeaux et la centrale météo (située dans la cabane). La centrale météo servant de point d'accès à l'ensemble du système. Il y a donc en tout cinq centrales de mesures : quatre centrales "micro-dendromètres" situées au niveau de chaque placeau et une centrale "météo" située dans la cabane (Figure 2.4).

Cette phase a été essentielle pour une planification précise de la préparation et l'installation du dispositif. Les différentes contraintes liées au site d'étude : envoi d'une partie du matériel par bateau, durée fixe de la mission, éloignement du site d'étude, construction des infrastructures rendaient cette planification d'autant plus importante.

### 2.3.2 Mise en place du dispositif

Un travail important a été réalisé en amont de l'installation du dispositif avec l'échantillonnage des arbres instrumentés ainsi que la préparation du terrain et du matériel de mesure. Ensuite, la mise en place du dispositif de suivi des micro-variations du rayon a été effectuée durant le mois de mai 2011 et a mobilisé les cinq personnes citées précédemment.

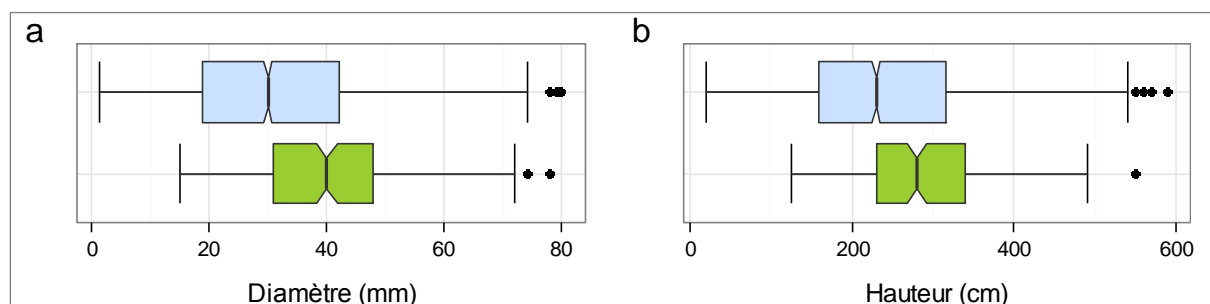


Figure 2.5 : Comparaison de la distribution entre l'échantillon des arbres suivis (vert) et la totalité de l'essai (bleu) pour le diamètre (a) et la hauteur (b) à 10 mois.



Figure 2.6 : Photo de la mise en place des infrastructures : a : cabane de gardien et local de rangement, b : installation des poteaux pour la clôture, c : câbles d'alimentations électrique et de communication entre centrales disposés pour être enterrés, d : fosse pour les mesures d'humidité du sol à différentes profondeurs.





### 2.3.2.1 Préparation

#### **Echantillonnage des arbres suivis**

Un inventaire de la hauteur et du diamètre a été réalisé 3 mois après la plantation de l'essai (février 2011), afin de présélectionner les arbres en fonction de leur représentativité et de leur localisation sur la parcelle. Cet échantillonnage ne présentait pas d'écart significatif par rapport à la globalité de l'essai en termes de moyenne et d'écart type ( $4,9 \pm 2,16$  mm de diamètre pour l'ensemble de l'essai et  $4,86 \pm 2,01$  pour l'échantillonnage initial). Fin avril 2011, la sélection définitive des arbres suivis a été effectuée après une deuxième mesure de diamètre. L'installation du support de capteur nécessitant un perçage du tronc, un diamètre minimal de 12 mm a été choisi pour sélectionner les arbres suivis. Cette contrainte technique a induit un biais dans l'échantillonnage : les arbres suivis sont en moyenne significativement plus gros que l'ensemble de l'essai (Figure 2.5). La variabilité captée dans l'échantillon est aussi plus faible que celle de la population totale mais reste importante avec un coefficient de variation phénotypique de 0,32 pour le diamètre et de 0,29 pour la hauteur à dix mois contre 0,41 et 0,37 pour l'ensemble de l'essai.

#### **Préparation du terrain**

La cabane de gardien ainsi que la clôture ont été mises en place début mai 2011, après un nettoyage de la parcelle fin avril 2011 (Figure 2.6). La clôture a nécessité 140 poteaux traités et 300 m de grillage. Le mât (3 m) de la station météo a été installé sur un socle en béton à côté de la cabane. Ce mât est situé à 20 m de la parcelle et à 6 m de la cabane pour éviter un effet de bordure lié à la proximité des arbres sur les mesures météo. Une tranchée de 40 cm de profondeur, allant de la cabane jusqu'à l'intérieur de la clôture a été creusée afin d'enterrer les câbles d'alimentation et de communication. Sur chaque plateau, une fosse de deux mètres de profondeur a été creusée pour permettre l'installation de quatre sondes d'humidité du sol. Deux poteaux par plateau ont été positionnés pour fixer les boîtiers contenant les centrales d'acquisitions de données et les multiplexeurs.

#### **Préparation du matériel**

Une partie de la préparation du matériel a été effectuée à Bordeaux. Le support de micro-dendromètre a été adapté pour prendre en compte la croissance rapide des eucalyptus. Des tests en conditions réelles de mesure ont été faits sur pin pour valider le prototype, puis 270 pièces ont été usinées (Annexe 1).

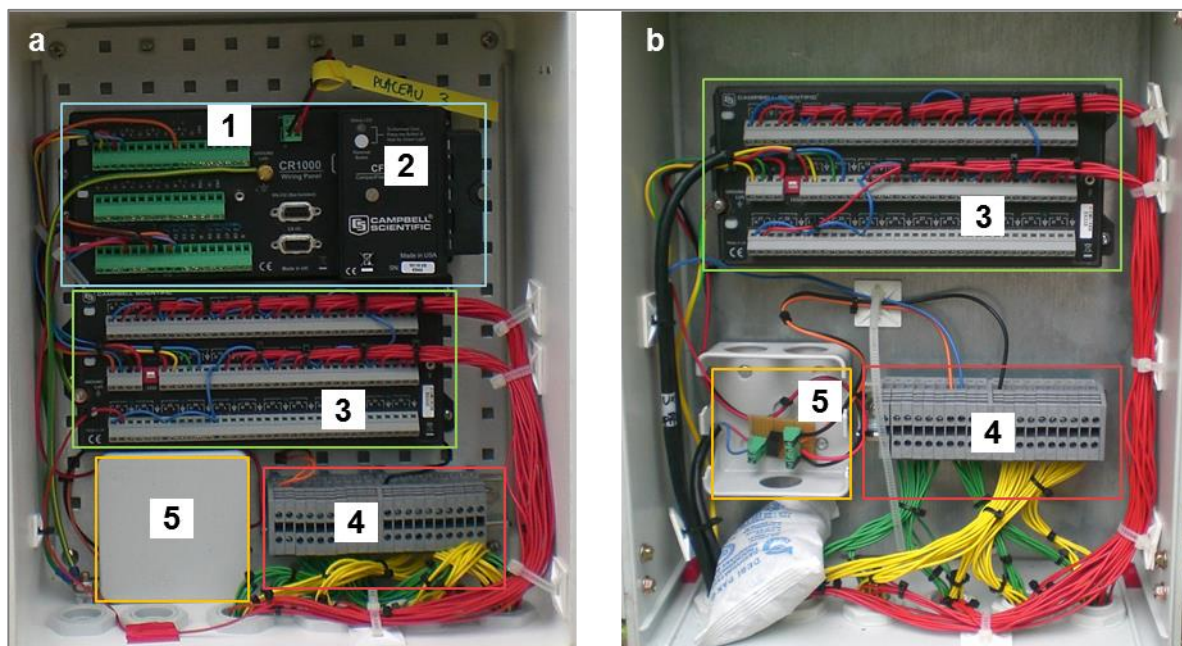


Figure 2.7 : Vue intérieure d'un boîtier "centrale de mesure" avant câblage final (a) et d'un boîtier multiplexeur (b). Centrale de mesure (1), module d'enregistrement (2), multiplexeur (3), bornier de connexion (4) et régulateur de tension 12,5V vers 2,5V (5).

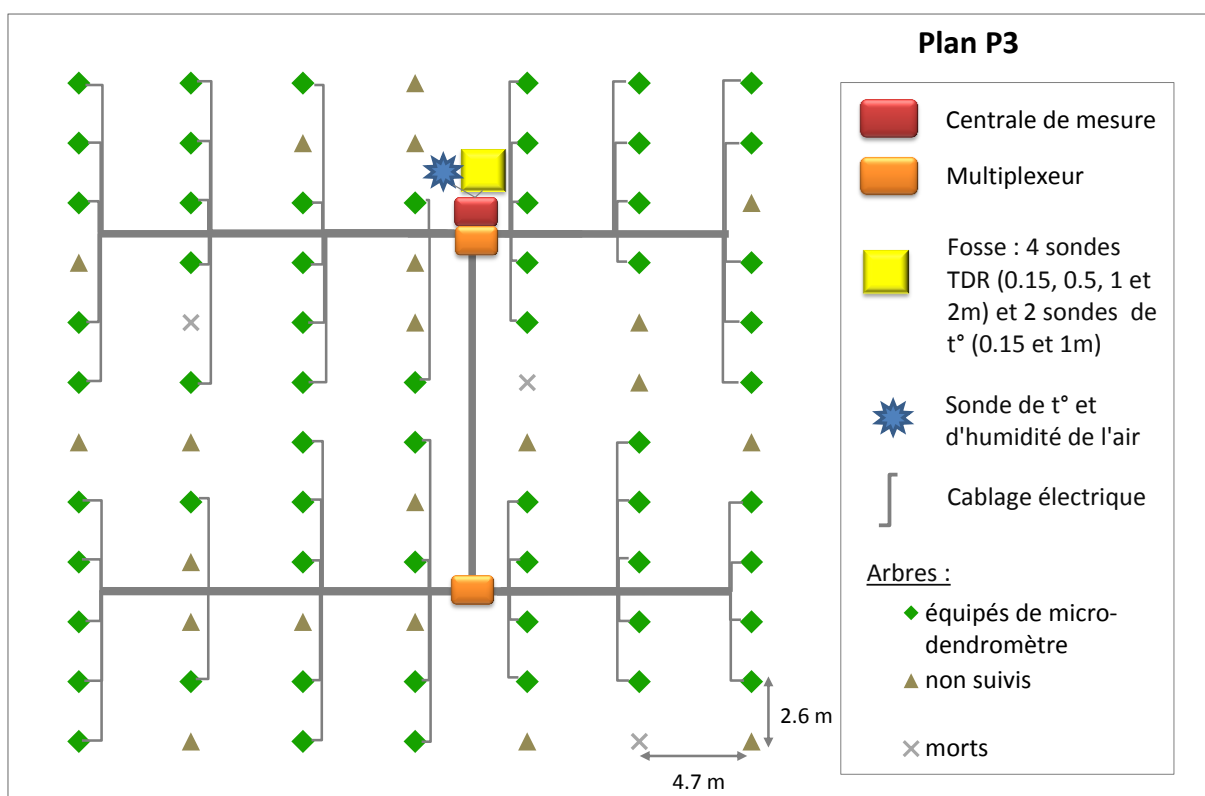


Figure 2.8 : Plan détaillé du plateau P3.



Les capteurs étant alimentés en 2,5 V, des circuits spécifiques convertissant du 12,5V (CC entrant) en 2,5V (CC sortant) ont été préparés pour être installés en amont de chaque multiplexeur. Un travail important de préparation des câbles a été réalisé à Pointe-Noire en avril 2011 (au total 4,5 km de câbles ont été nécessaires à ce dispositif). Afin de relier les capteurs à la centrale, des longueurs prédéfinies de câbles ont été découpées, puis les extrémités ont été dénudées pour pouvoir effectuer les branchements (Annexe 2). Deux boîtiers par plateau ont été préparés : l'un contenant la centrale de mesure et un multiplexeur, l'autre contenant le deuxième multiplexeur afin qu'il puisse être déporté de la centrale (Figure 2.7). Les branchements entre les centrales, les multiplexeurs et les capteurs ont été réalisés suivant un schéma précis. Chaque arbre suivi est associé à un capteur numéroté (avec un étalonnage spécifique), une longueur de câble et une voie de mesure sur le multiplexeur. La connexion entre la centrale et le capteur se fait par trois fils : rouge, vert et jaune (Figure 2.7).

#### *2.3.2.2 Description détaillée des éléments du dispositif*

A titre d'exemple, la Figure 2.8 représente de manière détaillée la structure du plateau 3. Un premier multiplexeur est relié à 31 arbres, le second à 29 arbres et le tout est connecté à centrale d'acquisition. Les sondes de température et d'humidité sont reliées directement à la centrale d'acquisition. Les autres plateaux sont équipés avec les mêmes composants mais diffèrent au niveau du nombre d'arbres suivis (P1 : 58 arbres, P2 : 64 arbres et P4 : 58 arbres). Les différentes mesures (micro-variations du rayon, variables environnementales) s'effectuent via cinq centrales (CR1000, Campbell Scientific). L'enregistrement des données s'effectue sur la mémoire interne des centrales ainsi que sur une carte mémoire, grâce à un module complémentaire d'enregistrement (CFM100, Campbell Scientific). Les programmes de mesure d'une centrale micro-dendromètre et de la centrale météo sont donnés en Annexe 3.

#### **Mesure des micro-variations du rayon**

La mesure des micro-variations du rayon de chaque arbre est réalisée toute les 30 secondes et moyennée toutes les cinq minutes pour être enregistrée. La mesure de tension variable, spécifique à chaque capteur est effectuée au niveau du multiplexeur, et la mesure de tension "totale" est effectuée de manière globale pour l'ensemble des capteurs d'un même multiplexeur (les capteurs sont branchés en parallèle). Le ratio entre ces deux tensions est calculé et enregistré : c'est le signal de micro-variation brut.

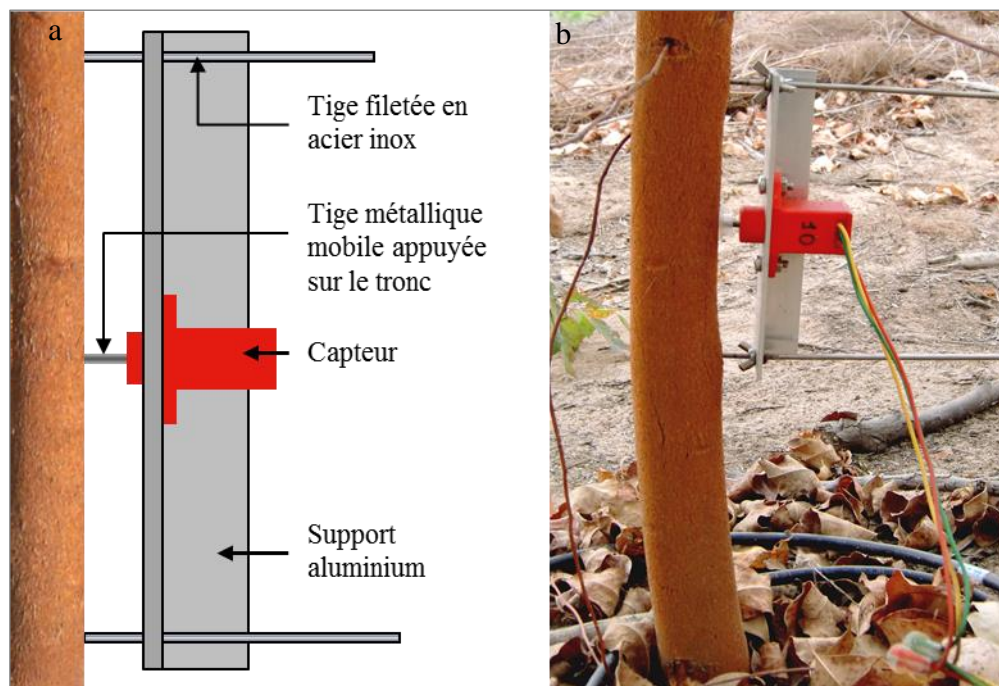


Figure 2.9: Vue schématique (a) et photo (b) d'un micro-dendromètre et du support aluminium.



Le capteur utilisé est un potentiomètre linéaire (résistance variable à trois bornes). Il est fixé à l'arbre par deux tiges en acier inoxydable et un support en aluminium maintenu par des écrous papillon (Figure 2.9). Tous les capteurs sont positionnés au sud. La connexion entre le capteur et la centrale d'acquisition est réalisée avec un câble blindé composé de trois fils (0,2 mm<sup>2</sup>). Deux capteurs par plateau sont placés sur un support inerte (tube en acier) pour servir de témoin. Pour mesurer l'hétérogénéité sur un même arbre, un arbre par plateau comporte trois capteurs. Les deux capteurs supplémentaires sont situés au même niveau mais avec une orientation différente (Est et Ouest).

### **Caractérisation environnementale**

Le suivi des variables environnementales se fait d'une part au niveau de la centrale météo pour la température, l'humidité de l'air, le rayonnement incident, la pluviométrie et la vitesse du vent et d'autre part au niveau de chaque plateau pour la température, l'humidité de l'air et l'humidité du sol. Les mesures sont enregistrées toutes les cinq minutes, au même pas de temps que les mesures de micro-variations. La mesure de température et d'humidité de l'air est effectuée sous abri par une sonde unique (CS215, Campbell Scientific). Le rayonnement incident est mesuré à trois mètres du sol par un pyranomètre (CS300, Campbell Scientific) avec une gamme spectrale de 300 à 1000 nanomètres. La pluviométrie est mesurée par un pluviomètre à auget basculant (0,2 mm/basculement, ARG100, Campbell Scientific) positionné à 1,3 m du sol. La vitesse et la direction du vent sont mesurées grâce à un anémomètre à coupelle (03002, Campbell Scientific) positionné à 3 m du sol avec une précision de 0,5 m/s (de 0 à 50 m/s). La température du sol est mesurée à deux profondeurs : 0,15 et 1 m (107, Campbell Scientific) avec une précision de 0,1°C. La teneur en eau du sol est mesurée par des sondes réflectométriques (CS616, Campbell Scientific) avec une précision de 0,05% (% volumique). Ces sondes sont placées à différentes profondeurs : 0,15 ; 0,5 ; 1 et 2 m afin de couvrir la fraction du profil présentant les plus fortes variations d'humidité. En effet, au-delà de deux mètres de profondeur, ces variations sont proches de ce que l'on observe à deux mètres [18].

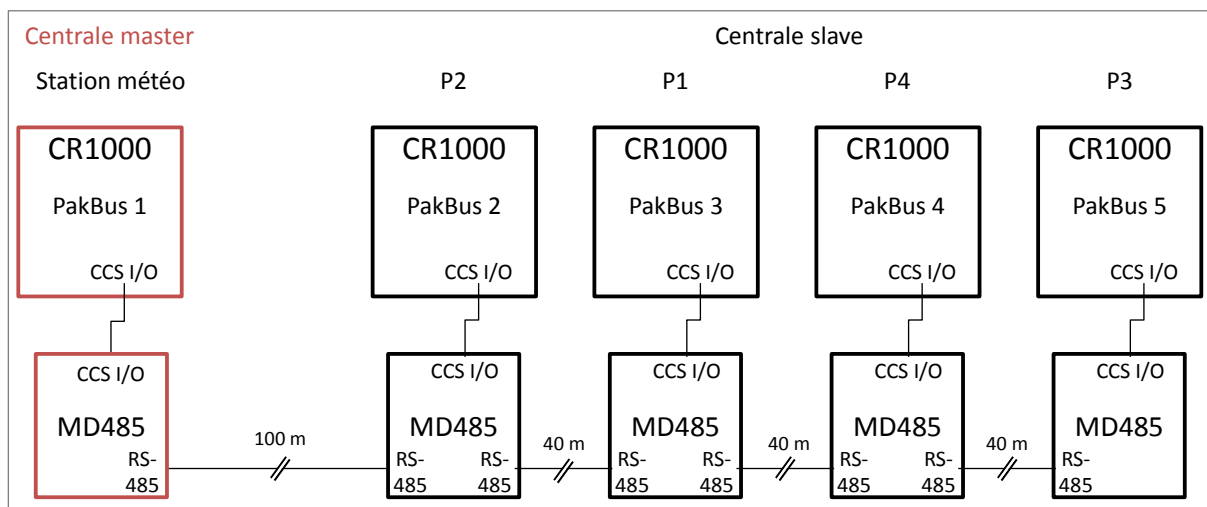


Figure 2.10 : Schéma du système de communication entre centrales (CR1000) via le module MD485. La centrale météo sert de centrale "master", les autres sont des centrales "slave".

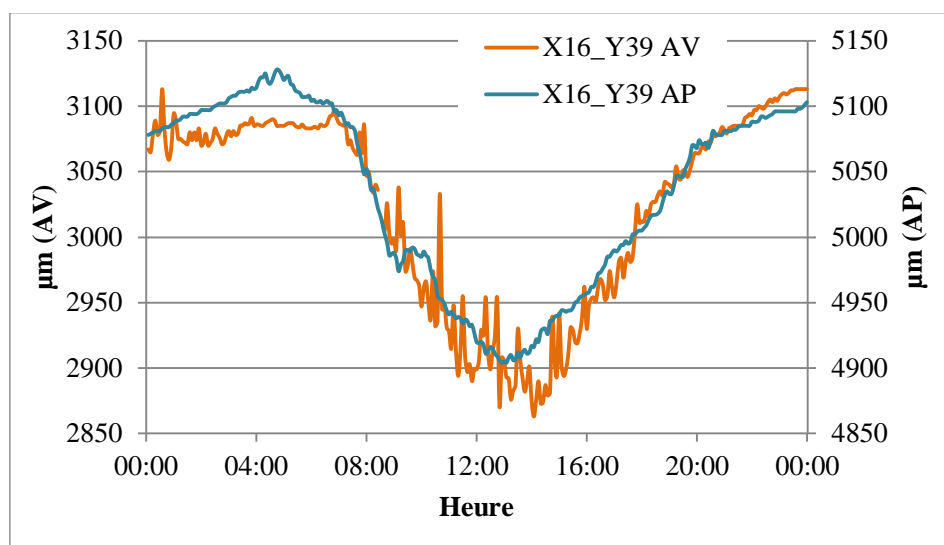


Figure 2.11 : Patron journalier de micro-variation du rayon sur l'arbre X16\_Y39 à deux périodes différentes : AV : sans régulateurs de tension et AP : avec régulateurs de tension. Les mesures représentent une journée complète (24h) pour les deux périodes.





### **Alimentation électrique et communications entre centrales météo et micro-dendromètres**

L'alimentation électrique est séparée en deux systèmes indépendants et identiques : le premier alimente les placeaux 1 et 4 et le deuxième alimente les placeaux 2 et 3 (Figure 2.4). L'alimentation électrique de chaque système est permise grâce à deux panneaux solaires (A-10P, ATERSA) branchés en parallèle, qui rechargent via un régulateur de tension (Solsum 6.6F, Steca Elektronik) deux batteries (12V, 140Ah) branchées en parallèle. Les deux batteries sont reliées aux deux placeaux via un câble électrique (2,5 mm<sup>2</sup>).

La communication entre centrales d'acquisition est opérée grâce un réseau (RS-485) organisé de la manière suivante. La centrale météo est la centrale "master" et les centrales micro-dendromètre sont des centrales "slave" (Figure 2.10). Les centrales sont reliées via le port CS I/O à un module complémentaire (MD485, Campbell Scientific). Les cinq modules sont reliés entre eux par un réseau filaire (2,5 mm<sup>2</sup>) qui permet la communication en RS-485. Les centrales ainsi que les modules MD485 sont configurés avec le même débit (115,2 kBd). Cette communication entre centrales permet l'accès à l'ensemble du système depuis la cabane de gardiennage via la centrale météo (Figure 2.4).

#### **2.3.3 Amélioration et suivi du dispositif Eucalyptus**

Des améliorations et des réparations ont été faites de manière ponctuelle lors de trois missions que j'ai effectuées à Pointe-Noire. Des régulateurs de tensions 12V/1,3A (THN 15-1212, Tracopower) ont été placés en amont de chaque centrale, afin de réduire les fluctuations de tension d'entrée liées aux panneaux solaires. Dans un même objectif de régulation de la tension du système, des convertisseurs de tension du 12V vers du 2,5V/1A (TSRN 1-2425A, Tracopower) ont été installés en amont de chaque multiplexeur. Cette régulation de la tension à deux niveaux a permis une forte diminution du bruit électrique présent dans le signal de micro-variations (Figure 2.11)

Des essais de transmission des données via le réseau de téléphonie mobile ont également été effectués afin d'envoyer les données automatiquement. La centrale météo munie d'un modem devait collecter les fichiers de données de chaque placeau, puis faire un envoi via le réseau mobile. Des difficultés liées à l'automatisation de la collecte de données et la mauvaise réception du réseau, combinées à la difficulté de faire des tests en n'étant pas sur place, n'ont

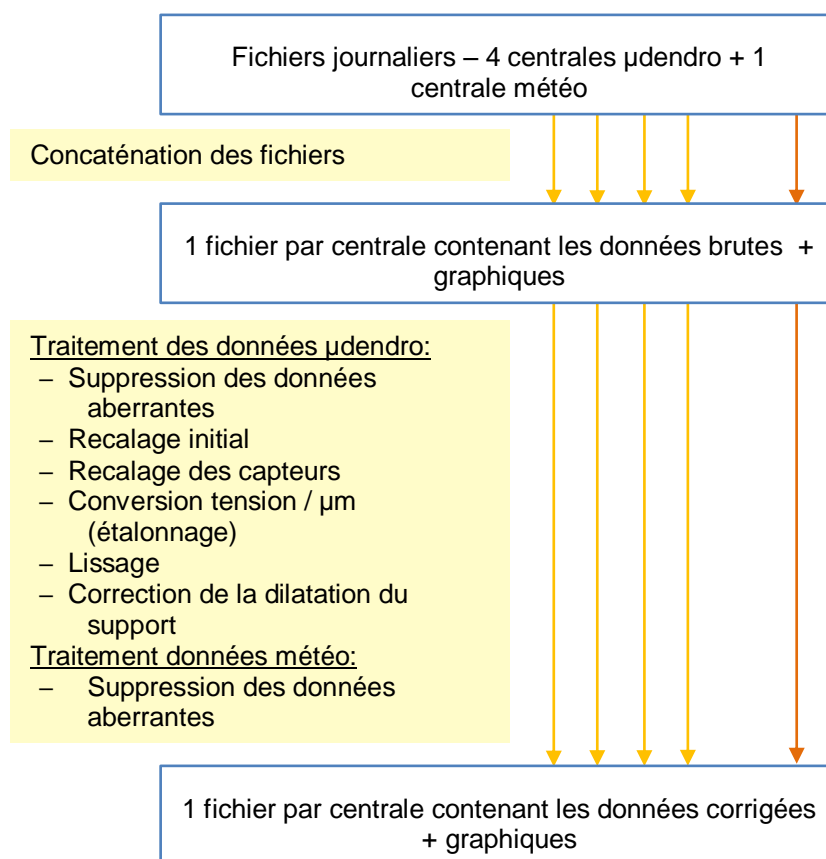


Figure 2.12 : Etapes de traitements des données brutes issues du dispositif P10. Les fichiers journaliers sont assemblés par centrales puis les données sont corrigées. Différents graphiques sont générés pour chaque variable : un graphique comprenant l'ensemble des données et des graphiques par période de deux semaines.

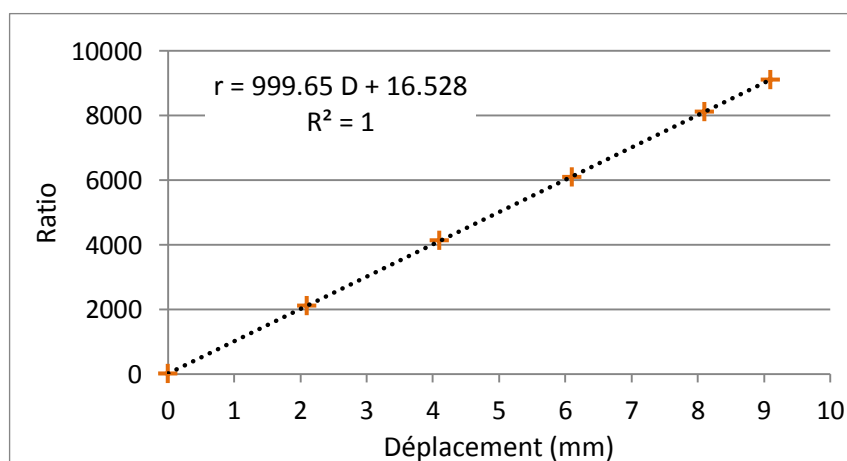


Figure 2.13 : Relation linéaire entre le déplacement et le ratio de tension aux bornes d'un micro-dendromètre.





pas permis de réaliser cet objectif. La récupération des données a donc été effectuée manuellement via la centrale météo grâce au logiciel dédié (Loggernet, Campbell Scientific).

Le suivi du dispositif a été assuré par André Mabiala (CRDPI) de façon hebdomadaire : récupération et envoi des données, recalage des capteurs de micro-variations et entretien des capteurs météo. Les capteurs de micro-variations nécessitent un suivi très fréquent. En effet, la course maximale du capteur étant de 10 mm cela impose un recalage du capteur avant d'arriver en bout de course. Dans notre cas, la croissance rapide des hybrides *E. urophylla* x *E. grandis* a imposé des recalages tous les deux mois en moyenne la première année, puis à une fréquence moindre. Le support a donc été pensé pour faciliter ce recalage et une cale de 8 mm d'épaisseur a été fabriquée pour servir d'étalon lors des recalages.

#### 2.3.4 Traitement des données

Le traitement des données environnementales et du signal de micro-variations a été effectué avec le logiciel R [19]. Les différentes étapes de traitements sont représentées dans la Figure 2.12. A chaque étape, des graphiques sont générés pour chaque variable afin de visualiser la qualité des données et du traitement effectué.

Les centrales génèrent un fichier par jour contenant 288 valeurs pour chaque variable mesurée, soit une mesure toute les cinq minutes. Le traitement de ces données commence par la concaténation de ces fichiers journaliers, qui permet de générer un fichier par centrale avec l'ensemble des données. Les données environnementales ( $T^\circ$ , HR, PP...) sont directement exploitables une fois les valeurs aberrantes éliminées. Par contre, les données de micro-variations nécessitent un traitement informatique plus important avant de pouvoir être utilisées. Les valeurs aberrantes sont tout d'abord supprimées. La mesure de tension (ratio) est ensuite convertie en déplacement ( $\mu\text{m}$ ) en utilisant l'équation linéaire propre à chaque capteur (exemple Figure 2.13). En effet chaque capteur a été étalonné indépendamment. Ensuite, la course initiale du capteur lors de sa mise en place sur l'arbre étant arbitraire, la première valeur de la période étudiée pour chaque capteur a été soustraite à l'ensemble des valeurs (recalage à 0 initial). Au cours de la croissance de l'arbre le capteur a été recalé plusieurs fois, faisant ainsi apparaître des sauts de valeurs (Figure 2.14, a). Ces sauts de valeurs ont été corrigés en additionnant la dernière valeur avant recalage et en soustrayant la première valeur après recalage à l'ensemble des données suivant le saut et ceci de manière itérative. On obtient ainsi la trajectoire de croissance radiale globale (Figure 2.14, b).

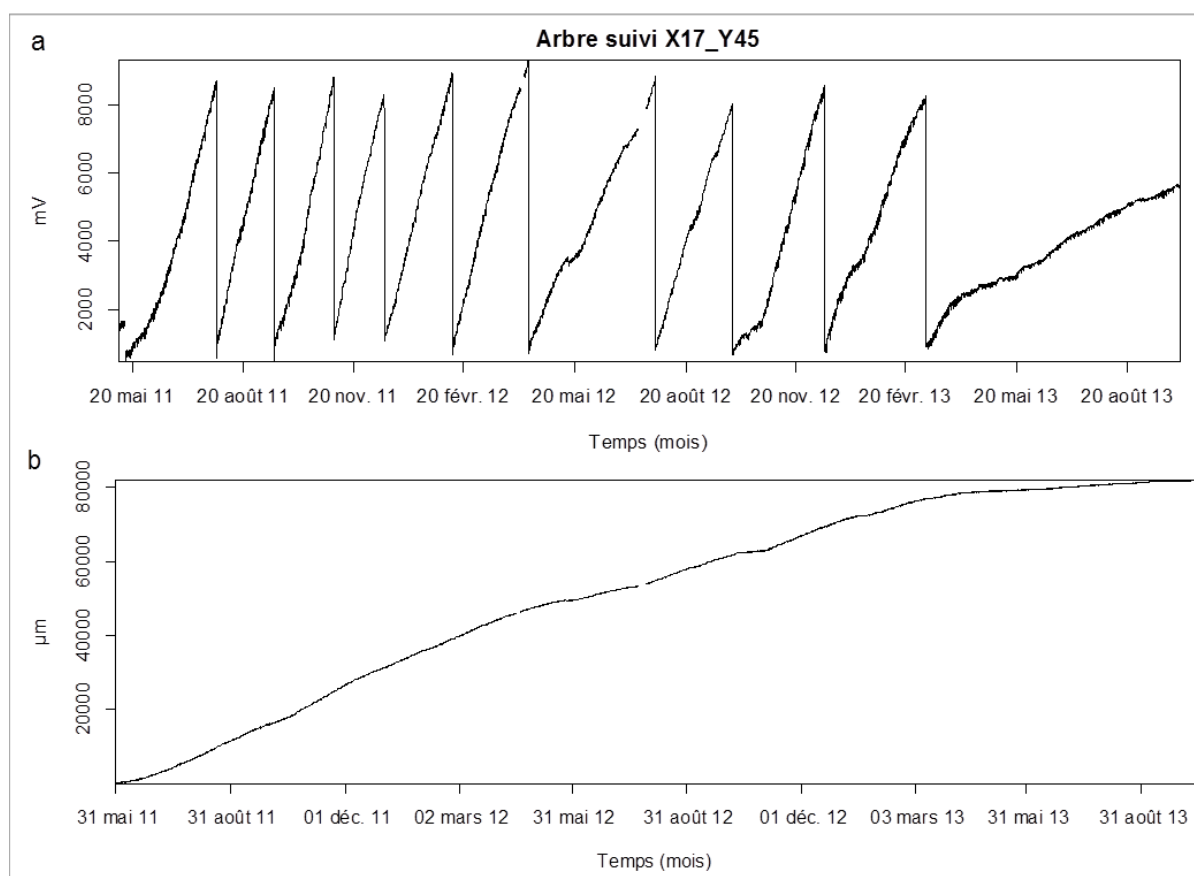


Figure 2.14 : Ensemble des données de micro-variations pour l'arbre X17\_Y45 : brutes (a) et après corrections (b).

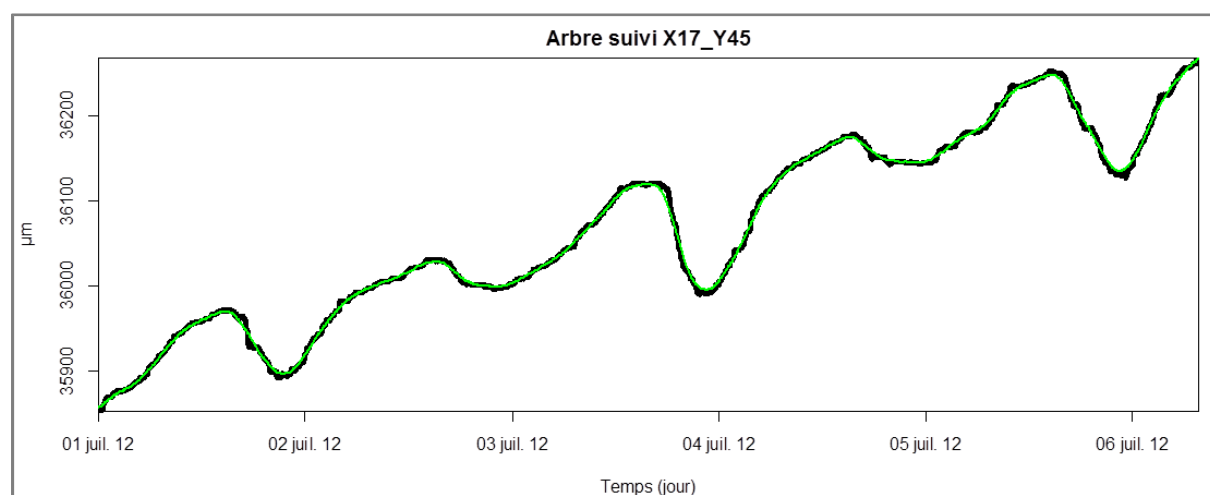


Figure 2.15 : Effet du lissage (courbe en vert) sur les données de micro-variations de l'arbre X17\_Y45.



Ces données sont ensuite lissées avec une moyenne mobile pour éliminer le bruit du signal (Figure 2.15). Enfin, l'effet de la température sur la dilatation des matériaux du support mais aussi sur la résistance globale du système est pris en compte grâce aux capteurs témoins sur support inerte. L'amplitude thermique moyenne durant la journée est relativement faible ( $8^{\circ}\text{C} \pm 2.2^{\circ}\text{C}$ ) sur la période de mesure. Une fois ces traitements effectués, les données corrigées sont exploitables pour l'analyse.

## 2.4 Caractères phénotypiques mesurés

### 2.4.1 Croissance

L'ensemble des essais a été mesuré pour des caractères ponctuels de croissance : hauteur et circonférence/diamètre. La hauteur a été mesurée à la perche au jeune âge puis à l'aide d'un dendromètre (Vertex IV, Haglöl, Suède). La mesure de circonférence a été réalisée avec un mètre ruban. Le diamètre a été mesuré avec un compas forestier. Les mesures de circonférence ont été faites à 1,3 m pour P93 et P97, et à 20 cm (hauteur du capteur) et 1,3 m pour P10. La fréquence des mesures est variable selon les essais. P93 a été mesuré à 14, 26, 39, 51 et 59 mois. Pour P97, 17 mesures sont disponibles : 7, 8, 9, 10, 11, 12, 13, 15, 18, 21, 24, 27, 30, 34, 44, 52 et 62 mois. P98 a été mesuré à 6, 12, 15, 25, 28, 33, 36, 42, 48, et 60 mois. Pour ces trois dispositifs antérieurs à la thèse (P93, P97 et P98) nous avons une vision globale de la croissance jusqu'à l'âge de rotation (5 à 7 ans). En revanche, les mesures ne sont disponibles que jusqu'à trois ans pour P10, qui a été planté au cours de la thèse : 3, 6, 10, 15, 18, 23, 27, 30 et 35 mois.

La croissance a également été modélisée afin d'obtenir la trajectoire globale (hauteur et croissance) pour les essais P93, P97 et P98. Le modèle monomoléculaire est un cas particulier du modèle de croissance von Bertalanffy-Richards [20, 21] qui a été largement appliqué pour modéliser la croissance des arbres [22-26]. L'équation du modèle est la suivante :

$$l(t) = \text{Asym} \left[ 1 - e^{-\text{lrc}(t-c_0)} \right]$$

$l(t)$  est la croissance cumulée à l'instant  $t$ ;  $\text{Asym}$  est la croissance maximale (asymptote);  $\text{lrc}$  est le logarithme de la constante,  $c_0$  est l'âge théorique pour lequel  $l(t)$  est égal à zéro. La modélisation de la croissance a été réalisée avec le package R: *nlme* [27, 28].





### 2.4.2 Mesures de $\delta^{13}\text{C}$

La composition isotopique du carbone ( $\delta^{13}\text{C}$ ) d'un échantillon est définie par l'équation suivante :

$$\delta^{13}\text{C} = \left( \frac{{}^{13}\text{C}/{}^{12}\text{C}_{\text{échantillon}}}{{}^{13}\text{C}/{}^{12}\text{C}_{\text{PDB}}} - 1 \right) \times 1000$$

Elle représente l'écart entre le rapport des isotopes stables du carbone de l'échantillon et celui du standard international (Peedee belemnite, PDB). Cette composition isotopique est différente entre l'air et la matière organique du fait de la discrimination du  $^{13}\text{CO}_2$  durant la phase de transport et de fixation dans les plantes. Le  $\delta^{13}\text{C}$  est directement relié à la discrimination du  $^{13}\text{CO}_2$  ( $\Delta$ ) par la plante avec l'équation suivante [29]:

$$\Delta = \frac{\delta_{\text{air}} - \delta_{\text{plante}}}{1 + \delta_{\text{plante}}/1000}$$

Avec une composition isotopique de l'air ( $\delta_{\text{air}}$ ) d'environ -8‰. Farquhar *et al.* [30] ont mis en lien la discrimination isotopique du carbone par la plante ( $\Delta$ ) et l'efficacité d'utilisation de l'eau intrinsèque (water use efficiency,  $\text{WUE}_i$ ) qui est défini comme le rapport entre l'assimilation de  $\text{CO}_2$  ( $A$ ) et la conductance stomatique pour l'eau ( $g_s$ ) [31]. De fait la composition isotopique du carbone de la plante ( $\delta^{13}\text{C}$ ) est utilisée comme un indicateur intégratif et robuste de la  $\text{WUE}_i$ . Cette mesure indirecte de la  $\text{WUE}_i$  a notamment été utilisée en amélioration génétique des plantes [32, 33] grâce à la rapidité de la mesure du  $\delta^{13}\text{C}$  par spectrométrie isotopique de rapport de masse qui permet une caractérisation des phénotypes à grande échelle.

Les mesures de  $\delta^{13}\text{C}$  ont été faites sur un compartiment différent de l'arbre selon les essais. Pour P93 et P97, les mesures ont été faites sur l'holocellulose extraite d'une portion de rondelle de bois. Cette rondelle a été prélevée à mi-hauteur à 59 mois pour le dispositif P93 et à 1,5 m à 62 mois pour le dispositif P97. Après broyage de la portion de rondelle, l'holocellulose a été extraite par la méthode Jayme-Wise [34] et placée dans des capsules d'étain (1 mg). Pour P10, cinq feuilles de lumière ont été prélevées à deux périodes différentes : en fin de saison sèche (septembre 2011) et en fin de saison des pluies (mai 2012). Ces feuilles ont ensuite été broyées et la poudre a été placée dans des capsules d'étain (1 mg). L'analyse de  $\delta^{13}\text{C}$  a ensuite été effectuée à Nancy sur le Plateau Technique d'Ecologie Fonctionnelle. La composition isotopique a été mesurée par un analyseur élémentaire en flux continu (Carlo Erba NA 1500; CE Instruments) couplé à un spectromètre de masse à rapport isotopique (Thermo-Finnigan; Delta S).

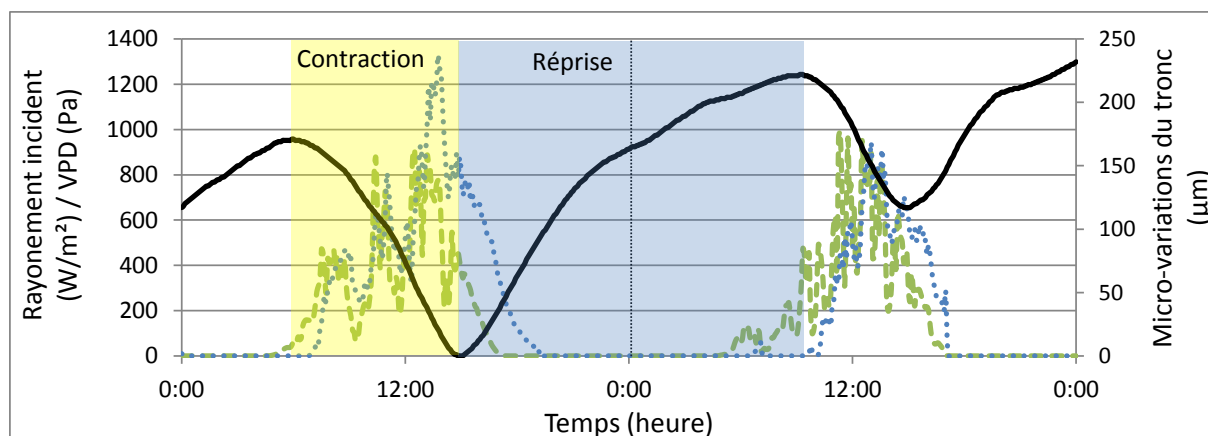


Figure 2.16 : Evolution des micro-variations du rayon (trait noir) en relation avec le rayonnement incident (tirets verts) et le VPD (pointillés bleus). La phase de contraction est en jaune et la phase de réhydratation en bleu.

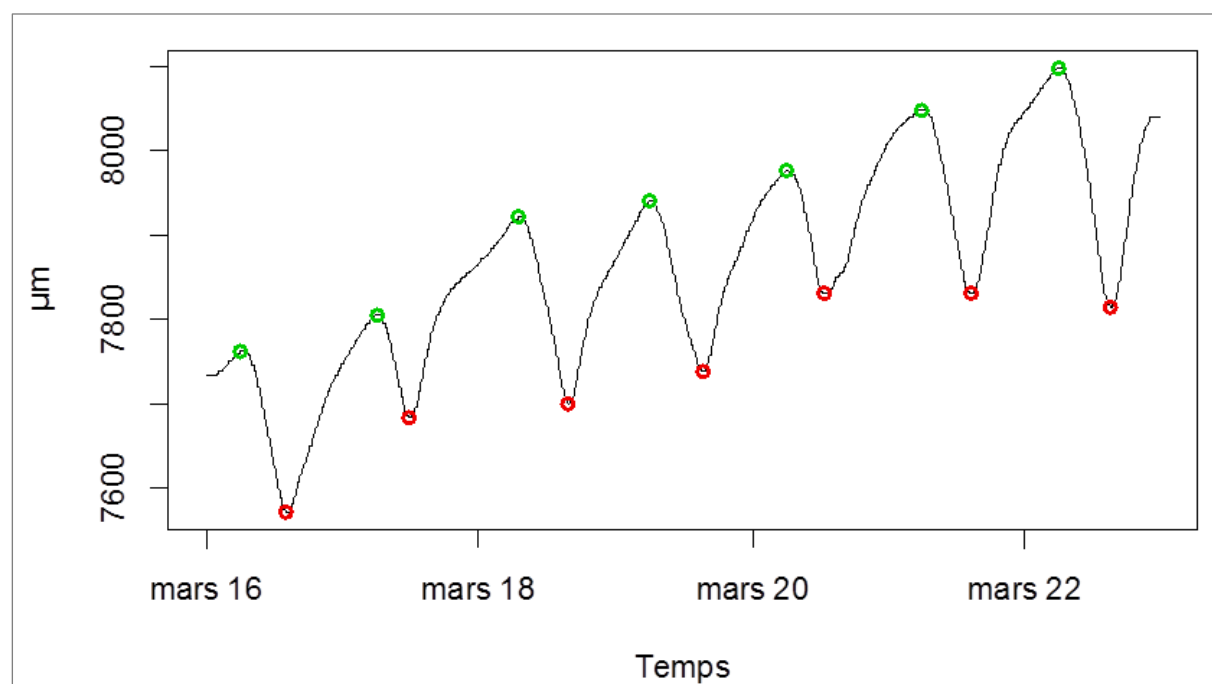


Figure 2.17 : Extraction du minimum (rouge) et du maximum (vert) journalier du signal de micro-variations du rayon.



### 2.4.3 Micro-variations du rayon

Les variations journalières du rayon du tronc ont été mises en relation avec le statut hydrique de l'arbre et les flux d'eau dans le tronc [15, 17, 35-39]. En effet, au cours de la journée, les réserves d'eau du tronc sont d'abord mobilisées vers les feuilles durant la phase de transpiration, ce qui induit une réduction du rayon du tronc (Figure 2.16). A l'inverse quand l'arbre ne transpire plus, une partie de l'eau absorbée au niveau racinaire est mise en réserve dans le tronc qui augmente donc de rayon (Figure 2.16). Ces micro-variations journalières, qui reflètent le contenu en eau du tronc sont donc très dépendantes de l'environnement (contenu en eau du sol d'un côté et demande évaporatoire de l'autre) [16, 40]. Elles sont principalement liées aux contractions des tissus élastiques : le bois en formation, le cambium et surtout le phloème [38, 41].

Le signal des micro-variations du rayon intègre également la croissance radiale (division du cambium, élongation cellulaire). Suivant l'échelle d'analyse, l'information apportée par le signal sur le statut hydrique et la croissance radiale n'est pas la même. Sur un pas de temps journalier, bien que de la croissance ait lieu, le signal traduit majoritairement le fonctionnement hydrique. Sur une échelle de temps plus longue (semaine, mois), il traduit essentiellement la croissance.

Les micro-variations du rayon ont été mesurées sur 240 arbres du dispositif P10. Après les étapes de traitements décrites précédemment, des variables ont été extraites du signal de micro-variation du rayon : le minimum et le maximum journalier (Figure 2.17). Ils permettent de calculer l'amplitude journalière, les accroissements entre maximums ou minimums successifs ainsi que les pentes associées aux phases de contraction ou de réhydratation.

## 2.5 Génotypage

Avec le développement des technologies de génotypage haut débit [42, 43], la diminution des coûts de génotypage permet à l'heure actuelle une augmentation considérable du nombre de marqueurs moléculaires et/ou d'individus génotypés. Dans ce contexte, il est nécessaire d'optimiser les étapes en amont comme l'extraction d'ADN.







### 2.5.1 Extraction d'ADN

Une nouvelle méthode plus rapide a été développée pour remplacer l'ancien protocole d'extraction d'ADN [44] peu adapté à l'extraction d'un grand nombre d'échantillons. L'extraction d'ADN de qualité à partir de feuilles d'eucalyptus est rendue difficile par la présence de composés métaboliques secondaires (polyphénols, terpènes) et de glucides. Différents protocoles ont donc été comparés et les meilleurs résultats en termes de qualité et de quantité ont été obtenus avec un protocole utilisant des billes magnétiques (Smart D-N-Adem-Kit). Ce protocole a ensuite été adapté pour l'extraction en plaque 96 puits (Annexe 4). Cette méthode d'extraction par billes magnétiques a pour avantage d'être automatisable. Les ADN utilisés pour le génotypage de la puce SNP ont été extraits en suivant ce protocole puis conservé à -20°C. La qualité et la quantité des ADN extraits ont été vérifiées par migration sur gel d'agarose (0,8%), spectrophotométrie (NanoDrop 8000) et spectrofluorimétrie (Quant-iT Picogreen).

### 2.5.2 Design de la puce SNP "Eucalyptus"

Le design de la puce SNP Infinium (Illumina, San Diego, USA) a été réalisé par Eric Mandrou au sein de la plateforme GénoToul à Toulouse. Tout d'abord les deux parents du croisement ont été entièrement séquencés (Illumina HiSeq2000) par la société GATC Biotech (Konstanz, Germany). Deux types de bibliothèques ont été générés par parent, l'une avec des fragments de 300 pb dont 100 pb séquencés aux deux extrémités, l'autre avec des fragments de 3kb séquencés sur 50 pb aux deux extrémités. Après alignement des séquences paires sur le génome de référence d'*E. grandis* (BRASUZ1, <http://www.phytozome.net/eucalyptus.php>), la détection du polymorphisme a permis de mettre en évidence plus de cinq millions de SNP chez les deux parents. Plusieurs étapes de sélection ont été appliquées afin d'obtenir 6000 SNP pour le design de la puce, principalement basées sur : i/ la robustesse du SNP (profondeur de séquençage, ADT score...), ii/ la couverture du génome (espacement entre SNP de 80 ou 300 kb) et iii/ la ségrégation attendue dans la descendance (1:1 ou 1:2:1).

Le génotypage de 1152 individus pour 6000 SNP a été réalisé par Genes Diffusion (Douai, France). Sur les 1152 individus génotypés, 48 sont les parents du croisement (24 répétitions de chaque), 60 sont issus du P93 (60 clones du P98) et 1044 sont issus du P10 (dont les 240 (dont les 240 arbres suivis en continu). Les résultats de génotypage ont été analysés avec le logiciel GenomeStudio (Genotyping module V1.9, Illumina).

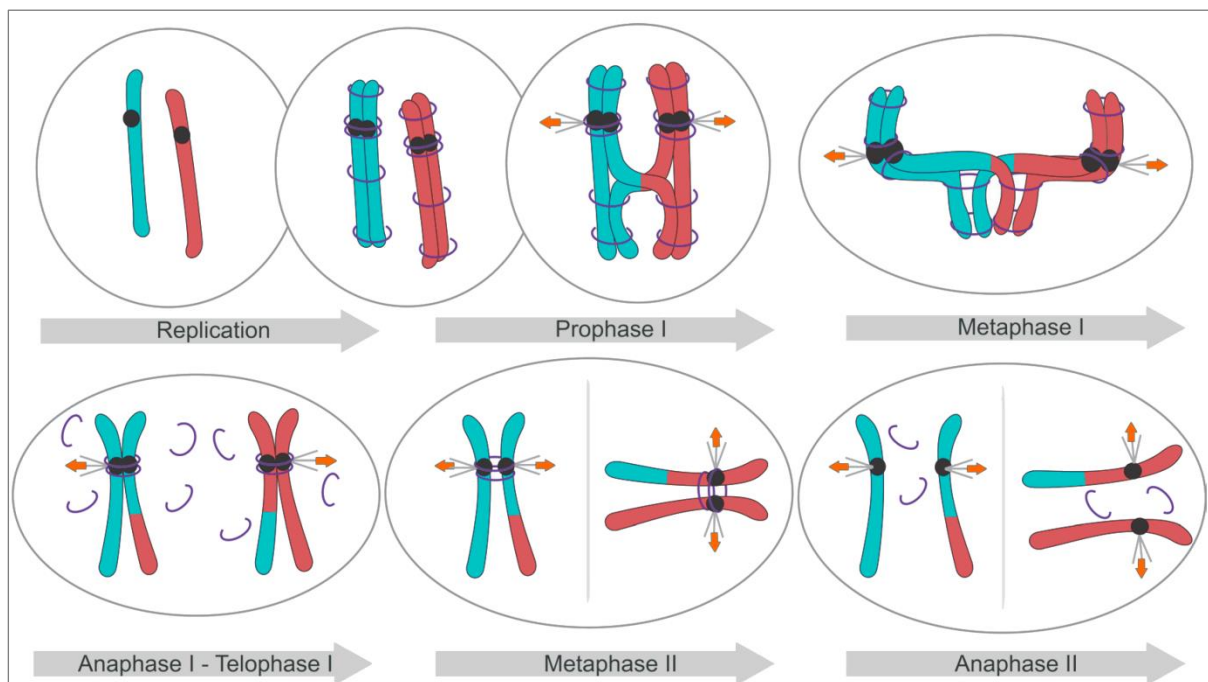


Figure 2.18 : Comportement schématique des chromosomes durant la méiose. La méiose est un type de division cellulaire qui aboutit à la formation de gamètes. Elle comprend deux phases de ségrégation des chromosomes après une seule réplication. Au début de la réplication, les chromatides sœurs sont maintenues ensemble (cercles violets). Ensuite au cours de prophase I, les chromosomes homologues s'apparient et le processus de recombinaison débute avec au moins un crossing-over par paire. Les crossing-over ainsi que la cohésion entre chromatides sœurs assure la liaison physique entre les chromosomes homologues recombinaisonnés, pour former un bivalent à la métaphase I. Durant l'anaphase I, cette cohésion est maintenue uniquement aux centromères: les chromosomes homologues se séparent vers les pôles opposés, tandis que les chromatides sœurs restent ensemble. À la métaphase II, la cohésion au niveau du centromère est rompue, ce qui permet, en anaphase II, la séparation des chromatides sœurs vers des pôles opposés. Schéma d'après Crismani *et al.* [45].



Les autres marqueurs moléculaires (RAPD, SSR et EST) disponibles pour les essais P93 et P97, sont antérieures à cette thèse.

## 2.6 Cartographie génétique

La cartographie génétique permet de représenter la répartition et l'ordre des locus le long des chromosomes. Elle s'appuie sur un phénomène majeur dans le maintien et la création de la diversité génétique [45, 46] : la recombinaison entre chromosomes homologues. Cette recombinaison se déroule lors de la formation des gamètes (méiose) durant la prophase I, avec la formation de crossing-over qui rompent le lien physique entre locus (Figure 2.18). Cet échange de portions de chromosome produit des haplotypes différents des haplotypes parentaux. En cartographie génétique, on cherche à estimer la fréquence d'occurrence des crossing-over le long du génome, grâce à l'observation de la "co-ségrégation" des locus dans une descendance.

On estime le taux de recombinaison ( $\theta$ ) entre deux locus (marqueurs moléculaire) par le ratio suivant :

$$\theta = \frac{\text{nombre de gamètes recombinés}}{\text{nombre de gamètes totaux}}$$

Le taux de recombinaison entre loci n'est pas une distance additive, il existe donc différentes fonctions [47, 48] qui permettent de le transformer en distance génétique (cM), afin d'obtenir la propriété d'additivité. La fonction de Kosambi, contrairement à celle d'Haldane, tient compte du phénomène d'interférence (réduction de la probabilité d'obtenir deux crossing-over proches l'un de l'autre). Cette différence se répercute au niveau des taux de recombinaison supérieurs 0,15 avec une augmentation de la distance estimée avec la fonction Haldane. Ces deux dernières restent les plus utilisées bien que d'autres fonctions aient été développées [49, 50].

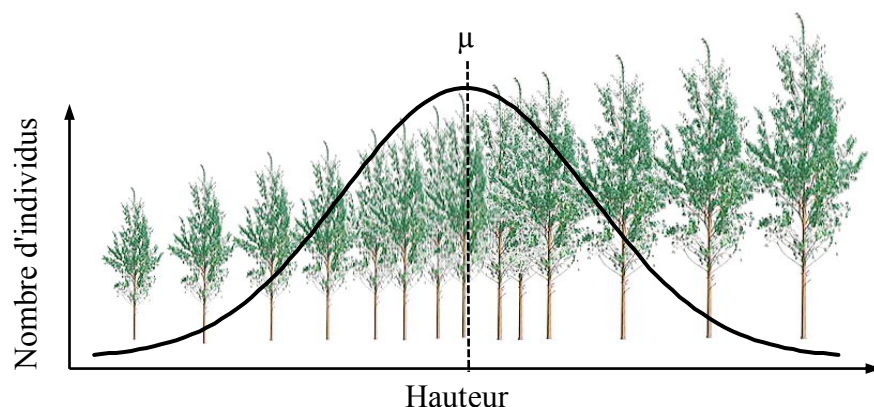
Différents facteurs influent sur la précision d'une carte génétique, comme le nombre d'individus, le type de population, le nombre de marqueurs moléculaires ainsi que leur type de ségrégation et l'algorithme de cartographie [51-54]. Dans le cas des arbres forestiers, les



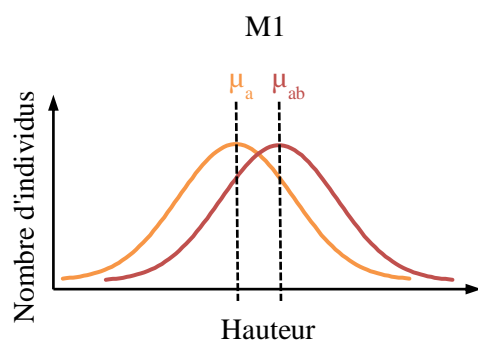
**Génotypage** : Mise en évidence du polymorphisme nucléotidique au niveau de différentes régions du génome (marqueurs moléculaire, M). Homozygote (aa) et hétérozygote (ab) .

Individu	1	à																							n
M1	aa	ab	aa	ab	aa	aa	aa	ab	ab	ab	ab	aa	aa	aa	ab	ab	aa	aa	ab	aa	aa	ab	aa	ab	...
M2	aa	aa	aa	ab	aa	ab	aa	aa	aa	ab	ab	ab	ab	ab	aa	aa	ab	ab	aa	aa	aa	ab	ab	aa	...

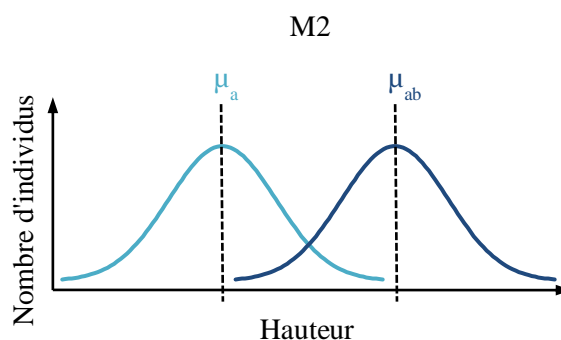
**Phénotypage** : Mesure d'un ou plusieurs caractères quantitatifs (ex: hauteur) dans une descendance. Le caractère doit présenter une distribution proche d'une distribution normale centrée sur la moyenne ( $\mu$ )



**Analyse QTL** : Test statistique entre la variation génétique au marqueur et la variation phénotypique



M1:  $\mu_a$  n'est pas significativement différente de  $\mu_{ab}$ . Il n'y a pas de QTL à proximité du marqueur M1



M2:  $\mu_a$  est significativement différente de  $\mu_{ab}$ . Il y a un QTL à proximité du marqueur M2

Figure 2.19: Principe de la détection de QTL



populations de cartographie sont issues d'individus fortement hétérozygotes du fait de la très faible domestication et du régime de fécondation préférentiellement allogame. Les possibilités en termes de types de populations sont donc limitées. En effet, les descendance F<sub>1</sub> et BC sont majoritairement utilisées, même s'il existe des études sur des familles F<sub>2</sub> [55]. La cartographie de ce type de population peut être effectuée avec la stratégie dite de "double pseudo-testcross" [56]. Cette stratégie permet de construire une carte génétique pour les deux parents, grâce à l'utilisation des marqueurs hétérozygotes chez un seul parent et homozygote chez l'autre, ségrégeant ainsi dans la descendance de pleins-frères dans les proportions mendélienne 1/2 : 1/2. La cartographie simultanée des deux parents (carte consensus) peut également être effectuée grâce à l'utilisation de marqueurs polymorphes chez les deux parents (SSR, SNP, DArT).

Dans cette étude nous avons utilisé deux logiciels de cartographie génétique qui permettent l'analyse simultanée de marqueurs avec différents types de ségrégation : JoinMap [57, 58] et le package R : *onemap* [59, 60]. JoinMap a été utilisé pour générer les cartes génétiques du P93 et P97. Celles du P10 ont été générées avec les deux logiciels. Cela a permis de comparer les résultats obtenus au niveau de l'ordre des marqueurs.

## 2.7 Détection de QTL

Un caractère qui présente une variation continue au sein d'une population (caractère quantitatif) est supposé être contrôlé par plusieurs gènes à effet faible : modèle infinitésimal [61]. Cependant, il a été montré que quelques gènes majeurs (QTL, quantitative trait locus) peuvent participer de façon significative à cette variabilité (modèle oligogénique). L'analyse QTL cherche à caractériser l'architecture génétique de ce type de caractère : c'est-à-dire à déterminer le nombre de régions du génome impliquées ainsi que leurs positions et leurs effets. Cette approche se base sur l'analyse combinée de l'information moléculaire et d'un caractère quantitatif dans une descendance en ségrégation. Elle permet de tester statistiquement le lien entre variation génétique (comme celle de marqueurs moléculaires) et variation phénotypique (Figure 2.19). Si ce test est significatif, on met en évidence un QTL.

Partant d'une analyse simple marqueur (analyse de variance à un facteur), l'analyse QTL a évolué vers des méthodes plus en plus complexes : i/ "l'interval mapping" (IM) [62, 63] utilisant l'information de la carte génétique pour positionner les QTL le long du génome ii/ le "composite interval mapping" (CIM), [64, 65] et iii/ le "multiple interval mapping" (MIM)



[66, 67], qui prennent compte des cofacteurs (marqueurs significativement associés au caractère) ce qui permet une meilleure estimation de la localisation et de l'effet des QTL. Parallèlement à ces méthodes classiques, des approches bayésiennes ont également été développées [68-70] mais restent pour l'instant peu utilisées chez les plantes.

Comme pour toutes les analyses statistiques, la taille de l'échantillon est un facteur critique. Pour des petites tailles d'échantillon, le risque de ne pas détecter de QTL à effet faible et donc de surestimer l'effet de ceux détectés est important. C'est ce que l'on appelle communément l'effet "Beavis" [71].

Dans cette étude, les différentes analyses QTL ont été réalisées avec le package R *qtl* [72, 73]. La méthode de CIM a été utilisée pour leur détection. Le taux d'erreur de type I au niveau du génome a été calculé avec dix mille permutations [74]. Les intervalles de confiance (95%) autour des QTL ont été calculés avec la fonction *bayesint*.

Les interactions QTL x E ont été testées en analyse de variance avec les fonctions *lm* et *anova* du logiciel R [19].



## Références

1. Gion J-M, Carouche A, Deweer S, Bedon F, Pichavant F, Charpentier J-P, Bailleres H, Rozenberg P, Carocha V, Ognouabi N *et al*: **Comprehensive genetic dissection of wood properties in a widely-grown tropical tree: Eucalyptus**. *BMC Genomics* 2011, **12**(1):301.
2. Gion J-M, Rech P, Grima-Pettenati J, Verhaegen D, Plomion C: **Mapping candidate genes in Eucalyptus with emphasis on lignification genes**. *Molecular Breeding* 2000, **6**(5):441-449.
3. Verhaegen D, Plomion C: **Genetic mapping in Eucalyptus urophylla and Eucalyptus grandis using RAPD markers**. *Genome* 1996, **39**(6):1051-1061.
4. Verhaegen D, Plomion C, Gion JM, Poitel M, Costa P, Kremer A: **Quantitative trait dissection analysis in Eucalyptus using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages**. *TAG Theoretical and Applied Genetics* 1997, **95**(4):597-608.
5. Steane DA, Nicolle D, McKinnon GE, Vaillancourt RE, Potts BM: **Higher-level relationships among the eucalypts are resolved by ITS-sequence data**. *Australian Systematic Botany* 2002, **15**(1):49-62.
6. Steane DA, Nicolle D, Sansaloni CP, Petroli CD, Carling J, Kilian A, Myburg AA, Grattapaglia D, Vaillancourt RE: **Population genetic analysis and phylogeny reconstruction in Eucalyptus (Myrtaceae) using high-throughput, genome-wide genotyping**. *Molecular Phylogenetics and Evolution* 2011, **59**(1):206-224.
7. Potts BM, Dungey HS: **Interspecific hybridization of Eucalyptus: key issues for breeders and geneticists**. *New For* 2004, **27**(2):115-138.
8. Corbasson M: **Eucalyptus urophylla**. *Revue Forestière Française* 1986, **38**:181-183.
9. Boland DJ, Brooker MIH, Chippendale G, Hall N, Hyland B, Johnston R, Kleinig D, McDonald M, Turner J: **Forest trees of Australia**, 5th edn. Collingwood: CSIRO publishing; 2006.
10. Vigneron P: **Création et amélioration de variétés hybrides d'Eucalyptus au Congo**. *'Bois et forêts des tropiques'* 1992, **234**:29-42.
11. Kottek M, Grieser J, rgen, Beck C, Rudolf B, Rubel F: **World Map of the Koppen-Geiger climate classification updated**. *Meteorologische Zeitschrift* 2006, **15**(3):259-263.
12. Mareschal L, Nzila JDD, Turpault MP, Thongo M'Bou A, Mazoumbou JC, Bouillet JP, Ranger J, Laclau JP: **Mineralogical and physico-chemical properties of Ferralic Arenosols derived from unconsolidated Plio-Pleistocene deposits in the coastal plains of Congo**. *Geoderma* 2011, **162**(1-2):159-170.
13. IUSS Working Group: **World reference base for soil resources 2006**. In: *World Soil Resources Reports*. ROME: FAO; 2006.
14. Impens II, Schalck JM: **A Very Sensitive Electric Dendrograph for Recording Radial Changes of a Tree**. *Ecology* 1965, **46**(1/2):183-184.
15. Garnier E, Berger A: **Effect of Water Stress on Stem Diameter Changes of Peach Trees Growing in the Field**. *J Appl Ecol* 1986, **23**(1):193-209.
16. Downes G, Beadle C, Worledge D: **Daily stem growth patterns in irrigated Eucalyptus globulus and E. nitens in relation to climate**. *Trees* 1999, **14**(2):102-111.
17. Simonneau t, Habib r, Goutouly j-p, Hugué j-g: **Diurnal Changes in Stem Diameter Depend Upon Variations in Water Content: Direct Evidence in Peach Trees**. *Journal of Experimental Botany* 1993, **44**(3):615-621.
18. Villar E: **Plasticité phénotypique et moléculaire de deux clones d'eucalyptus sous contrainte hydrique au champ**. *Ph.D.* Bordeaux: Université de Bordeaux I; 2011.
19. R Core Team: **R: A Language and Environment for Statistical Computing**. In: R Foundation for Statistical Computing, Vienna, Austria: URL <http://www.R-project.org/>; 2013.
20. von Bertalanffy L: **Quantitative Laws in Metabolism and Growth**. *The Quarterly Review of Biology* 1957, **32**(3):217-231.





21. Richards FJ: **A Flexible Growth Function for Empirical Use.** *Journal of Experimental Botany* 1959, **10**(2):290-301.
22. Pienaar LV, Turnbull KJ: **The Chapman-Richards Generalization of Von Bertalanffy's Growth Model for Basal Area Growth and Yield in Even - Aged Stands.** *Forest Science* 1973, **19**(1):2-22.
23. Zeide B: **Analysis of Growth Equations.** *Forest Science* 1993, **39**(3):594-616.
24. Vanclay JK: **Modelling forest growth and yield: applications to mixed tropical forests:** CAB International; 1994.
25. Fekedulegn D, Mac Siurtain MP, Colbert JJ: **Parameter estimation of nonlinear growth models in forestry.** *Silva Fennica* 1999, **33**(4):327-336.
26. Lei YC, Zhang SY: **Features and partial derivatives of Bertalanffy-Richards growth model in forestry.** *Nonlinear Analysis: Modelling and Control* 2004, **9**(1):65-73.
27. Pinheiro JC, Bates DM: **Mixed-Effects Models in S and S-Plus:** Springer; 2009.
28. Pinheiro J, Bates D, Debroy S, Sarkar D, the R Development Core Team: **nlme: Linear and Nonlinear Mixed Effects Models.** In., R package version 3.1-110 edn; 2013.
29. Farquhar G, Richards R: **Isotopic Composition of Plant Carbon Correlates With Water-Use Efficiency of Wheat Genotypes.** *Functional Plant Biology* 1984, **11**(6):539-552.
30. Farquhar G, O'Leary M, Berry J: **On the Relationship Between Carbon Isotope Discrimination and the Intercellular Carbon Dioxide Concentration in Leaves.** *Functional Plant Biology* 1982, **9**(2):121-137.
31. Farquhar G, Hubick K, Condon A, Richards R: **Carbon Isotope Fractionation and Plant Water-Use Efficiency.** In: *Stable Isotopes in Ecological Research.* Edited by Rundel PW, Ehleringer JR, Nagy KA, vol. 68: Springer New York; 1989: 21-40.
32. Condon A, Richards R, Rebetzke G, Farquhar G: **Breeding for high water-use efficiency.** *Journal of Experimental Botany* 2004, **55**(407):2447-2460.
33. Condon A, Richards R, Rebetzke G, Farquhar G: **Improving Intrinsic Water-Use Efficiency and Crop Yield Presented at the 1999 CSSA Symposium on Water Use Efficiency, organized by Div. C-2 chair, Dr. Tom Gerik.** *Crop Sci* 2002, **42**(1):122-131.
34. Leavitt SW, Danzer SR: **Method for batch processing small wood samples to holocellulose for stable-carbon isotope analysis.** *Analytical Chemistry* 1993, **65**(1):87-89.
35. Herzog K, Häsler R, Thum R: **Diurnal changes in the radius of a subalpine Norway spruce stem: their relation to the sap flow and their use to estimate transpiration.** *Trees* 1995, **10**(2):94-101.
36. Zweifel R, Item H, Häsler R: **Link between diurnal stem radius changes and tree water relations.** *Tree Physiology* 2001, **21**(12-13):869-877.
37. Cochard H, Forestier S, Améglio T: **A new validation of the Scholander pressure chamber technique based on stem diameter variations.** *Journal of Experimental Botany* 2001, **52**(359):1361-1365.
38. Sevanto S, Vesala T, Perämäki M, Nikinmaa E: **Time lags for xylem and stem diameter variations in a Scots pine tree.** *Plant, Cell & Environment* 2002, **25**(8):1071-1077.
39. Zweifel R, Zimmermann L, Newbery DM: **Modeling tree water deficit from microclimate: an approach to quantifying drought stress.** *Tree Physiology* 2005, **25**(2):147-156.
40. Deslauriers A, Morin H, Urbinati C, Carrer M: **Daily weather response of balsam fir (*Abies balsamea* (L.) Mill.) stem radius increment from dendrometer analysis in the boreal forests of Québec (Canada).** *Trees* 2003, **17**(6):477-484.
41. Zweifel R, Item H, Häsler R: **Stem radius changes and their relation to stored water in stems of young Norway spruce trees.** *Trees* 2000, **15**(1):50-57.
42. Gupta PK, Rustgi S, Mir RR: **Array-based high-throughput DNA markers for crop improvement.** *Heredity* 2008, **101**(1):5-18.





43. Appleby N, Edwards D, Batley J: **New Technologies for Ultra-High Throughput Genotyping in Plants.** In: *Plant Genomics*. Edited by Gustafson JP, Langridge P, Somers DJ, vol. 513: Humana Press; 2009: 19-39.
44. Doyle J, Doyle J: **Isolation of plant DNA from fresh tissue.** *Focus* 1990, **12**:13-15.
45. Crismani W, Girard C, Mercier R: **Tinkering with meiosis.** *Journal of Experimental Botany* 2013, **64**(1):55-65.
46. Schuermann D, Molinier J, Fritsch O, Hohn B: **The dual nature of homologous recombination in plants.** *Trends in Genetics* 2005, **21**(3):172-181.
47. Haldane JBS: **The combination of linkage values, and the calculation of distance between the loci of linked factors.** *J Genet* 1919, **8**:299-309.
48. Kosambi DD: **The estimation of map distances from recombination values.** *Annals of Human Genetics* 1943, **12**(1):172-175.
49. Tan Y-D, Fornage M: **Mapping functions.** *Genetica* 2008, **133**(3):235-246.
50. Zhao H, Speed TP: **On Genetic Map Functions.** *Genetics* 1996, **142**(4):1369-1377.
51. Ritter E, Gebhardt C, Salamini F: **Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents.** *Genetics* 1990, **125**(3):645-654.
52. Ferreira A, da Silva MF, Silva L, Cruz CD: **Estimating the effects of population size and type on the accuracy of genetic maps.** *Genetics and Molecular Biology* 2006, **29**(1):187-192.
53. Collard B, Mace E, McPhail M, Wenzl P, Cakir M, Fox G, Poulsen D, Jordan D: **How accurate are the marker orders in crop linkage maps generated from large marker datasets?** *Crop and Pasture Science* 2009, **60**(4):362-372.
54. Wu J, Jenkins J, McCarty J, Lou X-Y: **Comparisons of four approximation algorithms for large-scale linkage map construction.** *Theor Appl Genet* 2011, **123**(4):649-655.
55. Kole C: **Genome Mapping and Molecular Breeding in Plants: Forest trees**, vol. 7. Berlin: Springer; 2007.
56. Grattapaglia D, Sederoff R: **Genetic Linkage Maps of Eucalyptus grandis and Eucalyptus urophylla Using a Pseudo-Testcross Mapping Strategy and RAPD Markers.** *Genetics* 1994, **137**(4):1121-1137.
57. Stam P: **Construction of integrated genetic linkage maps by means of a new computer package: Join Map.** *The Plant Journal* 1993, **3**(5):739-744.
58. Van Ooijen JW: **JoinMap 4.1, Software for the Calculation of Genetic Linkage Maps in Experimental Populations.** In: Wageningen, Netherlands: Kyazma BV; 2011.
59. Margarido GRA, Souza AP, Garcia AAF: **OneMap: software for genetic mapping in outcrossing species.** *Heredity* 2007, **144**(3):78-79.
60. Mollinari M, Margarido GRA, Vencovsky R, Garcia AAF: **Evaluation of algorithms used to order markers on genetic maps.** *Heredity* 2009, **103**(6):494-502.
61. Fisher RA: **The genetical theory of natural selection:** Oxford University Press; 1930.
62. Broman KW: **Review of statistical methods for QTL mapping in experimental crosses.** *Lab Animal* 2001, **30**(7):44-52.
63. Zou W, Zeng Z-B: **Statistical Methods for Mapping Multiple QTL.** *International Journal of Plant Genomics* 2008, **2008**.
64. Jansen RC, Stam P: **High resolution of quantitative traits into multiple loci via interval mapping.** *Genetics* 1994, **136**(4):1447-1455.
65. Zeng ZB: **Precision mapping of quantitative trait loci.** *Genetics* 1994, **136**(4):1457-1468.
66. Kao C-H, Zeng Z-B, Teasdale RD: **Multiple Interval Mapping for Quantitative Trait Loci.** *Genetics* 1999, **152**(3):1203-1216.



67. Zeng ZB, Kao CH, Basten CJ: **Estimating the genetic architecture of quantitative traits.** *Genet Res* 1999, **74**(3):279-289.
68. Yi N, Shriner D: **Advances in Bayesian multiple quantitative trait loci mapping in experimental crosses.** *Heredity* 2007, **100**(3):240-252.
69. Ball RD: **Bayesian Methods for Quantitative Trait Loci Mapping Based on Model Selection: Approximate Analysis Using the Bayesian Information Criterion.** *Genetics* 2001, **159**(3):1351-1364.
70. Yandell BS, Mehta T, Banerjee S, Shriner D, Venkataraman R, Moon JY, Neely WW, Wu H, von Smith R, Yi N: **R/qtlbim: QTL with Bayesian Interval Mapping in experimental crosses.** *Bioinformatics* 2007, **23**(5):641-643.
71. Beavis WD: **QTL analyses: power, precision, and accuracy.** *Molecular dissection of complex traits* 1998, **1998**:145-162.
72. Broman KW, Sen S: **A Guide to Qtl Mapping With R/Qtl:** Springer; 2011.
73. Broman KW, Wu H, Sen S, Churchill GA: **R/qtl: QTL mapping in experimental crosses.** *Bioinformatics* 2003, **19**(7):889-890.
74. Churchill GA, Doerge RW: **Empirical Threshold Values for Quantitative Trait Mapping.** *Genetics* 1994, **138**(3):963-971.



### **3 Plasticity of primary and secondary growth dynamics in *Eucalyptus* hybrids**

RESEARCH ARTICLE

Open Access

# Plasticity of primary and secondary growth dynamics in *Eucalyptus* hybrids: a quantitative genetics and QTL mapping perspective

Jérôme Bartholomé<sup>1,2</sup>, Frédéric Salmon<sup>3</sup>, Philippe Vigneron<sup>4,5</sup>, Jean-Marc Bouvet<sup>3</sup>, Christophe Plomion<sup>2</sup> and Jean-Marc Gion<sup>1,2\*</sup>

## Abstract

**Background:** The genetic basis of growth traits has been widely studied in forest trees. Quantitative trait locus (QTL) studies have highlighted the presence of both stable and unstable genomic regions accounting for biomass production with respect to tree age and genetic background, but results remain scarce regarding the interplay between QTLs and the environment. In this study, our main objective was to dissect the genetic architecture of the growth trajectory with emphasis on genotype x environment interaction by measuring primary and secondary growth covering intervals connected with environmental variations.

**Results:** Three different trials with the same family of *Eucalyptus urophylla* x *E. grandis* hybrids (with different genotypes) were planted in the Republic of Congo, corresponding to two QTL mapping experiments and one clonal test. Height and radial growths were monitored at regular intervals from the seedling stage to five years old. The correlation between growth increments and an aridity index revealed that growth before two years old ( $r = 0.5$ ;  $0.69$ ) was more responsive to changes in water availability than late growth ( $r = 0.39$ ;  $0.42$ ) for both height and circumference. We found a regular increase in heritability with time for cumulative growth for both height [ $0.06 - 0.33$ ] and circumference [ $0.06 - 0.38$ ]. Heritabilities for incremental growth were more heterogeneous over time even if ranges of variation were similar (height [ $0-0.31$ ]; circumference [ $0.19$  to  $0.48$ ]). Within the trials, QTL analysis revealed collocations between primary and secondary growth QTLs as well as between early growth increments and final growth QTLs. Between trials, few common QTLs were detected highlighting a strong environmental effect on the genetic architecture of growth, validated by significant QTL x E interactions.

**Conclusion:** These results suggest that early growth responses to water availability determine the genetic architecture of total growth at the mature stage and highlight the importance of considering growth as a composite trait (such as yields for annual plants) for a better understanding of its genetic bases.

**Keywords:** Growth, Growth modelling, Plasticity, QTL x E interaction, Heritability, Eucalyptus

## Background

Production of wood biomass is a mandatory trait in any forest tree breeding program regardless of the final use be it pulp and paper, energy, construction or engineered wood products. Understanding the contribution of genetic and environmental factors as well as their interaction in tree growth and adaptation is a prerequisite for

accelerating tree domestication to meet the increasing demand for wood [1]. In this context, fast growing trees such as *Eucalyptus* will play a major role, not only as wood supply, but also as a model system to decipher the determinism of growth. Indeed, eucalypts are cultivated worldwide on more than twenty million hectares and are the most planted hardwoods in the world [2]. This genus comprises about 700 species [3] distributed naturally over a wide range of pedoclimatic conditions. Within that diversity, a few species or inter-specific hybrids combining the adaptive capacities and fast growth rates

\* Correspondence: jean-marc.gion@cirad.fr

<sup>1</sup>CIRAD, UMR AGAP, F-33612 Cestas, France

<sup>2</sup>INRA, UMR BIOGECO, F-33612 Cestas, France

Full list of author information is available at the end of the article

of the parental species (e.g. *E. urophylla* x *E. grandis* [4,5]), are widely used in planted forest.

Wood biomass formation is a dynamic process resulting from a wide array of physiological processes [6] some of which (e.g. photosynthesis) are dependent on environmental variations. It is also a product of height and circumference growth, involving the activity of the shoot apical meristem [7] and the vascular cambium [8]. The dynamics of both primary and secondary growth have been extensively described and modelled up to the rotation age in many forest tree species [9-12] with the main objectives of predicting biomass production and optimizing silvicultural practices. The first models corresponded to yield models and with the improved understanding of plant/tree physiological functioning and its interplay with climatic factors, process or mechanistic models were subsequently developed taking into account carbon allocation, nutrient and water availability, and climate effects [13,14]. These models were often developed for a particular species [15,16] without explicitly taking into account the genetic diversity within species, while a genetic effect of the response to environmental variation has been described in *Eucalyptus* on different scales of growth analysis (day, month, year) [17,18].

The genetic determinism of tree growth and its trend over-time have been fairly well studied in most forest tree species of commercial interest [19-24]. As regards time trends, the Franklin model [25], suggesting three phases in variance components for growth traits (juvenile, mature and adult phases), was tested in different species [26-28]. In short rotation species with rapid initial growth such as in *Eucalyptus*, a more stable trend was usually observed up to the rotation age [23]. Similarly, the ratio between dominance and additive variances has been analysed over time, with a major share of additive variance in *Pinus taeda* [29] and more balanced in *Eucalyptus* [23]. Age-age genetic correlations were found to be generally high between successive measurements and tended to decrease when the time lag increased i.e. between juvenile and mature stages [27,30,31]. Overall, these studies reveal low to moderate heritability for growth traits, indicating a strong environmental effect on the determinism of growth. More recent studies have tried to link genetic of growth response to environmental variation [23,32,33]. In *Eucalyptus globulus*, Costa e Silva *et al.* [34] reported a genetic effect on radial growth in response to temperature, rainfall and solar radiation. A similar result was presented by a more recent study showing a genetic effect for susceptibility to drought damage in *E. globulus* [35]. Nevertheless, there is still a need to improve our understanding of genetic and environmental determinants of growth, through measurements covering intervals better connected with environmental variations.

An analysis of the genetic architecture of growth traits (i.e. the number, map location and effect of quantitative trait loci, QTL) can also provide a better understanding of the genetic basis of growth. Thus, over the past two decades, many studies have been undertaken to detect QTLs controlling part of the phenotypic variation of wood biomass in forest trees such as pines [36], poplar [37] and eucalyptus [38]. In *Eucalyptus*, several genomic regions have been detected harbouring QTLs with low to moderate effects on the phenotypic variation of growth traits [39-44]. These studies, although carried out with a relatively small number of growth measurement time points revealed QTL instability over time [40,41]. The functional mapping approach taking into account the growth trajectory in QTL analysis [45] showed similar results with early and late QTLs in poplar [46]. However, rather few studies connecting the genetic architecture of growth traits and different environmental conditions have been carried out [47-50]. In *Eucalyptus* Freeman *et al.* [51,52] as well as Teixeira *et al.* [53], highlighted major QTL x Environment interactions. All these studies conducted on an inter-annual scale are consistent with the idea that growth is an integrative trait involving a large number of ecophysiological [6] and molecular [54] processes on which environmental conditions can act to determine the final genetic architecture of growth. Therefore, understanding the interplay between early growth and abiotic factors (e.g. alternation between dry and wet seasons in the tropics) and how environmental variations impact the genetics of growth is a key issue, especially as part of *Eucalyptus* breeding programmes.

In this context, the objective of this study was to dissect the genetic basis of the growth trajectory and its interplay with seasonal variation in water availability. To that end, one interspecific cross of *Eucalyptus urophylla* x *E. grandis* was planted in three different field trials in the Republic of Congo where contrasting environmental conditions in term of water availability (rainy vs. dry seasons), were suitable for observing the effect of drought on growth. We used two complementary types of growth trait: integrative traits (cumulative growth, growth curve parameters) and responsive traits (sub-annual growth increments) in combination with climatic data to characterize growth response to changes in water availability and reveal its genetic component. This comprehensive dissection of growth enabled us to characterize the genetic determinism of primary (height) and secondary (circumference) growth using conventional quantitative genetics and QTL detection approaches.

## Methods

### Plant material and field experiments

A full-sib family generated from a single controlled cross between *Eucalyptus urophylla* (accession 14.144) x *E.*



*grandis* (accession 9.21) was used in this study in three environmental settings corresponding to three field experiments. The first two trials were planted using a single tree plot design with seedlings: in April 1993 with 201 genotypes (referred to as P93) and already used in previous QTL studies [40,55,56] and in April 1997 (P97) with 190 genotypes (different from P93). The third trial was planted in November 1998 (P98) and corresponded to a clonal test with 60 genotypes in common with P93 replicated by rooted-cuttings with 10 copies per genotype.

The three trials were established in the Republic of Congo, in the Pointe-Noire region (4.8° S, 12° E). P93 and P97 were planted at a density of 667 trees/ha, whereas P98 was planted at 800 trees/ha. To reduce border effects a triple border row was planted around all the plantations.

DNA extractions for the full-sib progenies were carried out using dry leaves according to Doyle and Doyle [57].

#### Environmental characterization

The trials were planted on Ferralic Arenosols (FAO classification) at the Kissoko forestry station (Pointe-Noire, Republic of Congo). This soil is considered to be poor due to a low water retention, a small cationic exchange capacity and a low organic matter content [58]. The climate in the Pointe-Noire region corresponds to a tropical savanna climate [59] characterized by a dry season from May to October with annual rainfall of about 1,200 mm. The high atmospheric humidity (85% on average) and high temperature (25°C) display slight seasonal variations, at 2% and 3°C, respectively. To compare the climate between trials and especially the length of dry period during the dry season, we used the De Martonne aridity index ( $I_{DM}$ ).  $I_{DM}$  was calculated monthly with the following formula [60]:

$$I_{DM} = 12 p / (t + 10)$$

where  $p$  is the total monthly rainfall and  $t$  the monthly mean temperature. According to De Martonne [60], a threshold of  $I_{DM} = 15$  was chosen as an upper limit to define the dry season. Climatic data (temperature and rainfall) were recorded at Pointe-Noire about 25 km from the field trials.

#### Phenotypic traits

##### Growth Measurements

Tree growth was assessed by measuring both cumulative height and circumference at 1.3 m for each tree, at different ages. P93 was measured at 14, 26, 39, 51 and 59 months old. P97 was measured more frequently over time at 7, 8, 9, 10, 11, 12, 13, 15, 18, 21, 24, 27, 30, 34, 44, 52 and 62 months old. P98 was measured at 6, 12, 15, 25, 28, 33, 36, 42, 48, and 60 months old. The net

growth increments between two time points were calculated using the following formulas:

$$\text{Trait } n_1 - n_2 = \text{Trait } n_2 - \text{Trait } n_1$$

Where Trait  $n$  is the cumulative height or circumference at “ $n$ ” months old.

##### Growth modelling

We used the monomolecular growth equation to model height and circumference growth in trials P93, P97 and P98. The monomolecular growth equation is a particular case of the von Bertalanffy-Richards growth model [61,62] which has been widely applied and compared for tree growth modelling [63-67]. For fast growing trees like eucalypts, the monomolecular model is well adapted because of its concave shape with no inflection point [68-70]. The equation is expressed in the following form:

$$l(t) = \text{Asym} \left[ 1 - e^{-\text{lrc}(t-c_0)} \right]$$

Where  $l(t)$  is the cumulative growth at time  $t$ ; Asym is the maximum growth (the asymptote); lrc is the logarithm of the growth constant and  $c_0$  is the theoretical age for which  $l(t)$  is zero. Growth modelling was performed with *R* software (Version 2.15.2; <http://cran.r-project.org/>) using the *R* package: *nlme* [71,72]. Trees with a high proportion of missing data were removed, based on the duration and the phase (exponential growth phase or asymptotic phase) at which missing data occurred. Selected trees presented on average 0.3, 2.2 and 0.5% of growth curve that is missing for P93, P97 and P98 respectively. For each selected genotype, a nonlinear model (monomolecular) was fitted to obtain the above-mentioned growth curve parameters (GCP) for both height and circumference. The model fitted the data well as illustrated for circumference in Additional file 1. Moreover the pseudo R-squared calculated for all adjusted models were higher than 0.9. Predicted values were also calculated for P97 for the same time points as in P93 (14, 26, 39, 51 and 59 months old).

##### Broad-sense heritability of growth

Broad-sense heritabilities were estimated using trial P98 which corresponded to a clonal test with 10 replicates of 60 full-sibs. Analyses were conducted using Markov Chain Monte Carlo methods implemented in the *R MCMCglmm* package [73]. The following mixed effects model was used for each trait:

$$y = \mu + \text{Row}_i + \text{Col}_j + G_k + \varepsilon$$

where  $y$  is the individual tree measurement,  $\mu$  is the general mean,  $\text{Row}_i$  is the fixed effect of the row,  $\text{Col}_j$  is the fixed effect of the column,  $G_k$  is the random effect of the genotype  $\sim N(0, \sigma_G^2)$  and  $\varepsilon$  is the residual random

effect  $\sim N(O, \sigma_\varepsilon^2)$ . The intersection between individual tree coordinates in the trial (row and column) represented the spatial position of each tree. Broad-sense heritability for the full-sib family ( $H^2$ ) was estimated for each trait using the following formula:

$$H^2 = \sigma_G^2 / (\sigma_G^2 + \sigma_\varepsilon^2)$$

Where  $\sigma_G^2$  is the genetic variance and  $(\sigma_G^2 + \sigma_\varepsilon^2)$  is the phenotypic variance [74].

#### Molecular markers and linkage mapping

For each parental genotype (*E. urophylla* and *E. grandis*) and each trial (P93 and P97), a genetic map was constructed according to a two-way pseudo-test cross mapping strategy [75]. For P93, the two genetic linkage maps were constructed using RAPD SSR, EST, and STS markers genotyped on 201 individuals [56]. For P97, only RAPD markers selected on the basis of their map position in the P93 linkage maps were used for genetic mapping with 190 individuals. The set of common polymorphic markers genotyped in both P93 and P97, enabled the unambiguous identification of homologous linkage groups (LGs) between the two parental maps of the same species. In addition, the codominant markers used in P93, enabled the identification of orthologous regions between *E. urophylla* and *E. grandis* [56].

JoinMap® version 4.1 [76,77] was used to perform the linkage analysis. Chi-squared tests were performed to test whether markers followed the expected Mendelian segregation ratios. Distorted markers ( $p < 0.01$ ) were excluded from linkage analysis. Markers with more than 40% of missing data were also excluded. For the four maps, markers were grouped into LGs at the independence logarithm-of-the-odds (LOD) score  $\geq 4.0$ . Then, markers within linkage groups were ordered using the Monte Carlo maximum likelihood mapping algorithm [78,79]. Finally, mapping quality tools were used to classify markers leading to a poorer fit as accessory markers, located near the closest framework markers. Accessory markers with significant deviation from expected Mendelian segregation ( $p < 0.01$ ) were indicated by \* on the genetic maps. Genetic distances in centiMorgan (cM) were calculated with the Kosambi mapping function [80]. The LG nomenclature corresponds to that defined by Brondani *et al.* [81].

#### QTL analysis

##### QTL mapping

Quantitative trait locus (QTL) analyses were performed with *R* software using the *qtl* package [82]. We carried out simple and composite interval mapping using the multiple imputation method of Sen and Churchill [83] with a step interval of 1 cM, a genotyping error rate of

0.001 and 512 imputations per genotypes. It has been shown that this imputation method performs better with missing genotypic data [83]. The whole-genome significance thresholds for all traits were determined with 10,000 permutations [84]. QTLs that met or exceeded 95<sup>th</sup> and 90<sup>th</sup> percentiles were assessed to be significant or suggestive, respectively. Confidence intervals for each QTL position were obtained using Bayes credible intervals (BCI). Genetic maps and QTL positions were drawn using MapChart [85].

#### Detection of QTL x environment interaction

For each parent, we were able to test the QTL x environment interaction (QTL x E) using markers common to both trials (P93 and P97), and for growth traits at five different ages (14, 26, 31, 51, 59 months old) based on measurements for P93 and predicted values at the same ages for P97. GCP were also used for both trials. To test the QTL x E, an analysis of variance (*lm* and *anova* function, *R* software) was carried out according to the following model:

$$y = \mu + \text{Trial} + \text{Marker} + \text{Trial} \times \text{Marker} + \varepsilon$$

where  $y$  is the individual tree measurement,  $\mu$  is the general mean, Trial is the fixed effect of the trial (P93 or P97), Marker is the fixed effect of the closest marker associated with a QTL detected by CIM, Trial  $\times$  Marker is the interaction between the two fixed effects and  $\varepsilon$  is the residual random effect  $\sim N(O, \sigma_\varepsilon^2)$ .

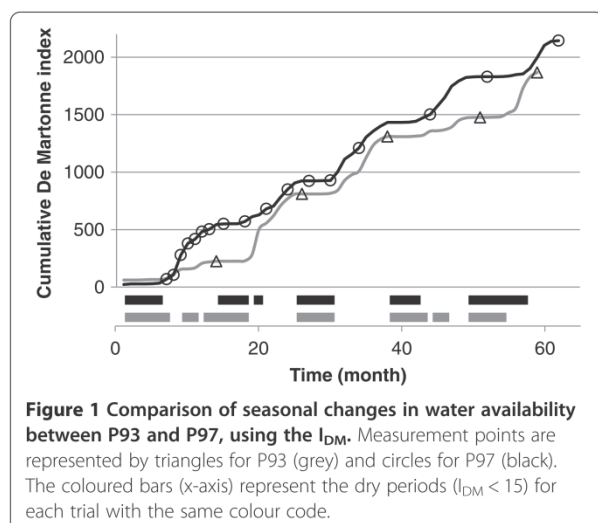
## Results

#### Climatic conditions

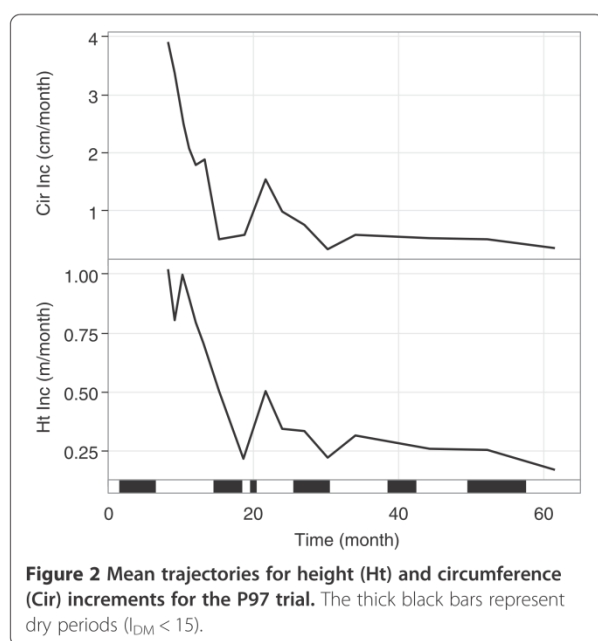
The De Martonne index ( $I_{DM}$ ) was calculated monthly during the growth period of P93 and P97 (Figure 1).  $I_{DM}$  mainly reflected seasonal changes in water availability because the mean temperature remained very stable across seasons with only 3°C of amplitude. The two curves exhibited different patterns at the beginning of the plantation (between 8 and 20 months), then evolved in parallel before separating again after 45 months. The timing and the intensity of the dry periods (defined as  $I_{DM} < 15$ ) were different between the two trials. Indeed, P93 exhibited longer dry periods than P97 (except for the third and the last dry seasons), amounting to four supplementary months of drought over the study period. Using the dry periods defined by  $I_{DM}$ , specific growth increments were determined for P97 only, because measurement points were lacking for P93.

#### Trend in phenotypic variation and correlation

The mean trajectories for cumulative growth, height (Ht) and circumference (Cir), were slightly different between the three trials (Additional file 2). Mean Cir values were



significantly smaller in P98 ( $p < 0.01$ ) whereas final Ht values at 60 months old were higher ( $p < 0.01$ ), probably due to a higher tree density in P98. These differences were reduced between P93 and P97, two trials planted at the same period (April) and with the same tree density. The mean trajectory of growth increments, estimated from P97 (Ht\_Inc and Cir\_Inc), showed a global decrease over the 60 months after planting, with two periods of lower growth rates that occurred during the second and third dry periods (Figure 2). These two periods flanked a growth rate peak at 21 months that occurred during the second rainy season.



The pattern of phenotypic variations for cumulative growth (Ht, Cir) was very similar between trials. Indeed the variance increased gradually over time and the coefficient of phenotypic variation (CVp) for Ht remained rather stable ([0.07-0.09], [0.13-0.19] and [0.12-0.24] for P93, P97 and P98, respectively, Additional file 3). The same tendency was observed for Cir. The CVp of growth increments for P97 and P98 were more variable compared to cumulative growth, as exemplified by CVp values for Ht\_Inc ranging from 0.18 to 0.58 for P97 and from 0.12 to 0.82 for P98. A similar range of values was observed for Cir\_Inc. Given the large number of measurements in P97, we were able to characterize the evolution of CVp over time for Ht\_Inc and Cir\_Inc. Two growth rate peaks CVp were observed during the second and the third dry periods (Additional file 4), corresponding to a lower growth increment on average. These results suggest that growth rate differences between genotypes were enhanced during these two dry periods.

This hypothesis was supported by the analysis of Pearson correlations between height or radial increments and  $I_{DM}$  calculated at both trial and tree levels using data from P97 (Table 1). At trial level (i.e. considering all data points for all trees), positive correlations were found for both early (<24 months) and late (>24 months) growth, indicating a strong effect of rainfall on growth. For both levels, early growth was found to be more responsive (higher correlations) than late growth, although fewer data were available at the later stage, which could have affected the estimation of correlation coefficients. The tree level reflected the phenotypic plasticity of each genotype in response to water availability (1 = highly responsive, non-significant = non responsive). At tree level, correlations between  $I_{DM}$  and circumference increments were higher than those with height increments (Figure 3). These differences were particularly apparent during early growth with 95% of trees having a coefficient of correlation over 0.5, including 81% of trees with a significant correlation ( $p$ -value < 0.05). This suggests that secondary growth was more prone to changes in water availability than primary growth, especially at the juvenile stage.

The analysis of phenotypic correlations between cumulative growth (Ht and Cir) revealed strong correlations (>0.78) over time in each trial. As expected, the age-age correlations for both cumulative growth traits (Ht and Cir) were highly positive and stable over time, and comparable between the three trials (Additional file 5; A,B,C). For Ht and Cir, the correlations between current age and final growth increased dramatically over the two first years to reach at least 0.8 at two years old (Additional file 6). High correlations were also found between Asym and the final growth. More interestingly, Pearson correlation coefficients for growth increments were variable over time. Moreover, correlations for Ht\_Inc and Cir\_Inc displayed roughly opposite patterns



**Table 1 Pearson's coefficients of correlation (r) between growth increments and  $I_{DM}$**

		Trial level	Tree level	
		Pearson's r (p-value)	Mean of Pearson's r [range]	% of significant correlation
Height increment	Early growth	0.5 (2.2e-16)	0.58 [-0.26, 0.96]	32.3%
	Late growth	0.39 (2.2e-16)	0.44 [-0.36, 0.9]	5.7%
Circumference increment	Early growth	0.69 (2.2e-16)	0.79 [0.02, 0.98]	81.1%
	Late growth	0.42 (2.2e-16)	0.49 [-0.08, 0.92]	7.5%

Pearson's coefficients of correlation were calculated for early (< 24 months) and late (>24 months) growth at trial and tree levels.

during early growth (until the third dry period). As illustrated for P97 in Figure 4, correlations for Ht\_Inc dropped from high and significant levels (maximum of 0.64 at 13 months old) to non-significant levels revealing contrasting behaviour between genotypes in their response to the environment between the beginning of the second dry period and the 34<sup>th</sup> month. Conversely, correlations for Cir\_Inc became significant during the same period revealing a similar response between genotypes.

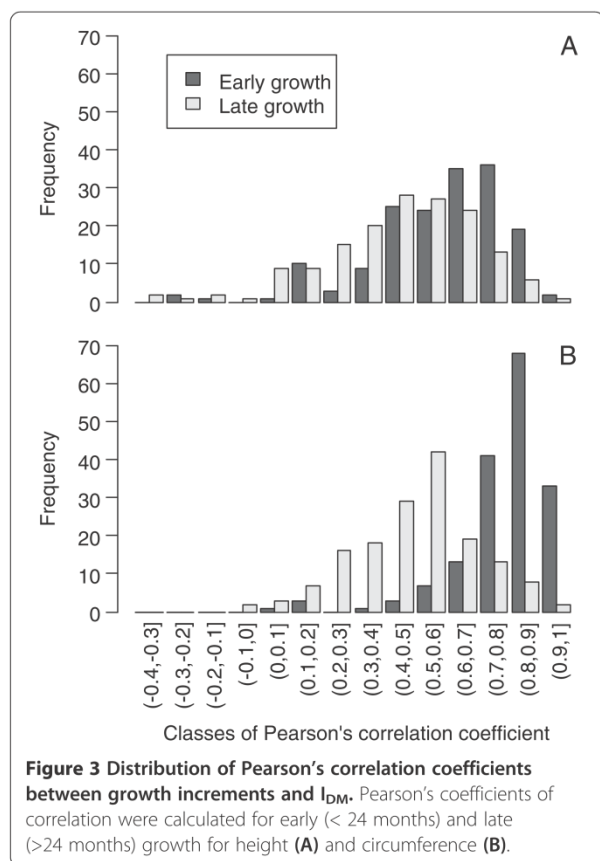
#### Trend in the genetic control of growth traits

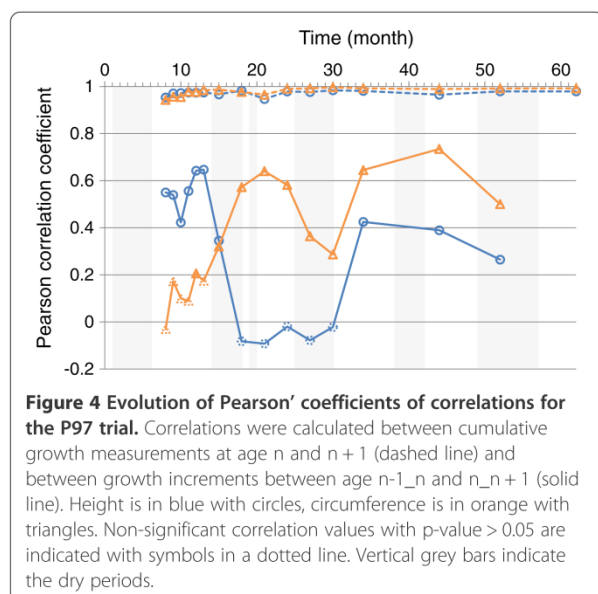
Broad-sense heritabilities ( $H^2$ ) were estimated for cumulative growth traits, growth increment, and the GCP

based on the data from P98. For all traits,  $H^2$  values were small to medium. Although standard errors associated with  $H^2$  were high ( $\pm 0.1$  on average), we observed a regular increase in heritability with time for cumulative growth: Ht [0.06-0.33], Cir [0.12 to 0.38] (Figure 5, Additional file 7). The heritabilities for the Asym parameter (Asym\_h: 0.35 and Asym\_c: 0.44) were only slightly higher than those obtained for final Ht and Cir, which is consistent with the strong correlation between these pairs of traits.  $H^2$  for growth increments were more heterogeneous than  $H^2$  for cumulative traits (Figure 5, Additional file 7), even though the variation ranges were similar (Ht\_Inc [0-0.31]; Cir\_Inc [0.19 to 0.48]). Given the standard error around  $H^2$ , only differences between cumulative and incremental growth for Ht (at ages 28, 33, 36 and 48) were found to be significant between the two types of trait, while no significant differences for Cir were observed.  $H^2$  for Ht\_Inc and Cir\_Inc had nearly an opposite pattern from 25 to 35 months, then their patterns became more similar. The lowest heritabilities for Ht and Cir clearly corresponded to the coefficient of variation peak occurring during the third dry season (Additional file 4) and suggest a greater environmental effect at that period.

#### Linkage mapping

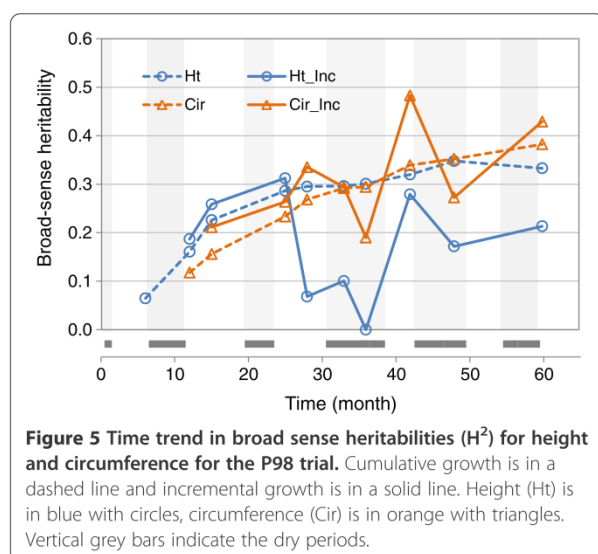
The linkage analysis was performed for the P93 and P97 trials separately. Table 2 summarizes the characteristics of each genetic map. For P93, 57 co-dominant markers mapped by Gion *et al.* (2011) enabled the identification of orthologous regions between *E. urophylla* and *E. grandis* parents. These regions covered a total of 989 cM for *E. urophylla* and 847 cM *E. grandis*, amounting respectively to 66% and 67% of the total map length. The macro-colinearity of common markers between parental maps was well conserved except for 6 small inversions in LGs 2, 5 and 7 (Additional file 8). For P97, the two parental maps were constructed in this study, using a total of 148 RAPD markers for *E. urophylla* and 130 RAPDs for *E. grandis*. A total of 106 and 86 framework markers were positioned in 13 (chromosome 2 and 11 were split into two sub-groups) and 11 linkage groups (LGs) for *E. urophylla* and *E. grandis* parents





respectively. The two maps provided a resolution of one marker every 14.5 cM (*E. urophylla*) and 15.3 cM (*E. grandis*) for a total map length of 1,179 cM (*E. urophylla*) and 1,021 cM (*E. grandis*).

Among the framework markers, 72 markers for the female map and 57 markers for the male map were common between P93 and P97 maps. These "bridge markers" enabled the identification 13 homologous regions for the *E. urophylla* parent covering 73% (P93) and 90% (P97) of total map length. For *E. grandis*, 11 homologous regions were identified, covering of 79% (P93) and 89% (P97) of the total map length. As expected, the order between markers was well conserved between P93 and P97: 89% and 93% of the common markers presented



the same order for the *E. urophylla* and the *E. grandis* maps, respectively. The four genetic maps with links between bridge markers revealing homologous and orthologous regions are represented in Additional file 8.

## Genetic architecture of growth traits

### QTL detection in P93

A total of 19 QTLs were detected for the female parent (*E. urophylla*) and 9 for the male parent (*E. grandis*) at the genome-wide type I error rate of 10% (Additional file 9). 71% of the QTLs remained significant at the 5% genome-wide error rate. These QTLs were associated with 15 and 8 different traits for *E. urophylla* and *E. grandis* respectively; representing 8 and 6 different regions distributed over 5 different linkage groups for each parental map. For the *E. urophylla* parent, the variance explained by all the detected QTLs associated with one trait ranged from 5.5% (Ht39) to 21.2% (Asym\_h) with an average of 11.5% ( $\pm 3.4$ ). For the male parent, this proportion of variance ranged from 4.2% (Ht39) to 13.2% (Cir39) with an average of 7.6% ( $\pm 3.1$ ). QTLs for *E. urophylla* were located on LG1, LG3, LG4, LG6 and LG7 with two hotspots (LG3 at 11-12 cM and LG6 at 22-27 cM) gathering 16% and 47% of the total number of QTLs respectively. QTLs for *E. grandis* were located on LG1, LG2, LG3, LG5 and LG10 with two hotspots (LG5 at 113-116 cM and LG10 at 27 cM) gathering 22% and 33% of the total number of QTLs, respectively. The distribution of QTLs between the two parents was different, of the seven LGs with QTLs only two (LG1 and LG3) were common to both parents. Two specific QTLs were found for the growth curve parameter on the *E. urophylla* map on LG3 (for Asym\_h and lrc\_c). No QTL for GCP were found on the *E. grandis* map.

### QTL detection in P97

Unlike P93, fewer QTLs were detected for the female parent (33) than for the male parent (72) at the genome-wide type I error rate of 10%, while 74% of the QTLs remained significant at 5%. These QTLs associated with 29 and 62 different traits, represented 8 and 9 different regions distributed on 6 and 7 different LGs for *E. urophylla* and *E. grandis*, respectively (Additional file 10). For *E. urophylla*, the variance explained by all the detected QTLs associated with one trait ranged from 5.1% (Cir11) to 15.3% (Cir52\_62) with an average of 7.6% ( $\pm 2.5$ ). This proportion of variance ranged from 5.1% (Cir8\_9) to 17% (Ht52) with an average of 10.5% ( $\pm 2.9$ ) for the male parent (*E. grandis*). QTLs for *E. urophylla* were located on LG2, LG3, LG5, LG6, LG8 and LG10 with three hotspots (LG5 at 47 cM, LG8 at 41-44 cM and LG8 at 66-68 cM) gathering 12%, 34% and 16% of the total number of QTLs, respectively. QTLs for *E. grandis* were located on LG2, LG3, LG4, LG5, LG6, LG8 and LG10 with three hotspots (LG2

**Table 2 Genetic map statistics for *E. urophylla* and *E. grandis* in both the P93 and P97 trials**

	<i>E. urophylla</i>		<i>E. grandis</i>	
	P93	P97	P93	P97
Number of LGs	11	13	11	11
Total map length (cM)	1,502	1,179	1,270	1,021
Average LG length (cM)	137	91	115	93
Number of framework markers	146	106	118	86
Mean distance between markers	11	15	12	15
Number of common framework markers between P93 and P97	72		57	
Missing data (%)	9.7	6.3	8.1	4.2

LG linkage group.

at 0-10 cM, LG4 at 73 cM and LG8 at 22-42 cM) gathering 6%, 6% and 60% of the total number of QTLs, respectively. All LGs with QTLs were common to both parents except for LG4, for which no QTL was detected in *E. urophylla*. One specific QTL region was found for the GCP located on the *E. grandis* map on LG10 (Asym\_c, lrc\_c and c0\_c). The other GCP QTLs co-located with QTLs for different traits.

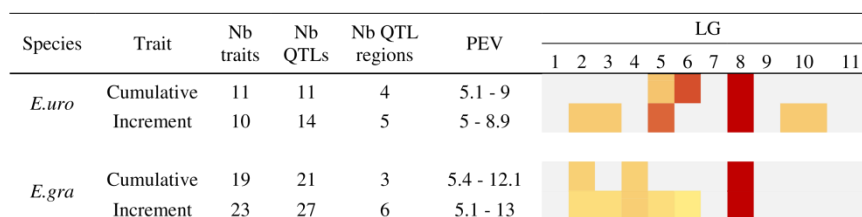
QTLs for growth increment were located on more LGs than QTLs for cumulative growth (Figure 6). The range of phenotypic variance explained by QTLs was similar between cumulative growth and growth increment. For both parents more significant QTLs (5% at genome wide level) were found for growth increments. These genomic regions were located on LG2, LG3 and LG10 for *E. urophylla* and on LG3, LG5 and LG6 for *E. grandis*. In addition to these specific QTLs for growth increments, one QTL region for *E. urophylla* on LG6 was found to be associated with early cumulative growth only (Cir 12, 13, 15, 18). More interestingly, all QTLs for late cumulative growth for both parents co-localized with QTLs for early growth increments, as illustrated for *E. urophylla* on LG5 at 47 cM and for *E. grandis* on LG2 at 0/10 cM and on LG8 at 22/42 cM (Figure 7). Moreover, a co-localization with a QTL for the Asym\_c parameter was found for two of these QTLs (*E. urophylla* LG2 and *E. grandis* LG8) suggesting that the genetic

architecture of growth response (incremental traits) at the early stage affected the genetic architecture of cumulative growth at the adult stage. Additionally, on LG2, a QTL for the correlation coefficient between growth increment and  $I_{DM}$  was found (Position: 0 cM, LOD = 2.55, BCI: 0-21 cM and PEV = 5.34%) for *E. grandis* (P97). This QTL co-located with the hotspot on LG2 for *E. grandis*. For this QTL region, genotypes with a higher growth were also the more responsive (higher Pearson correlation coefficient between Cir\_Inc and  $I_{DM}$ ).

#### QTL comparison across trials

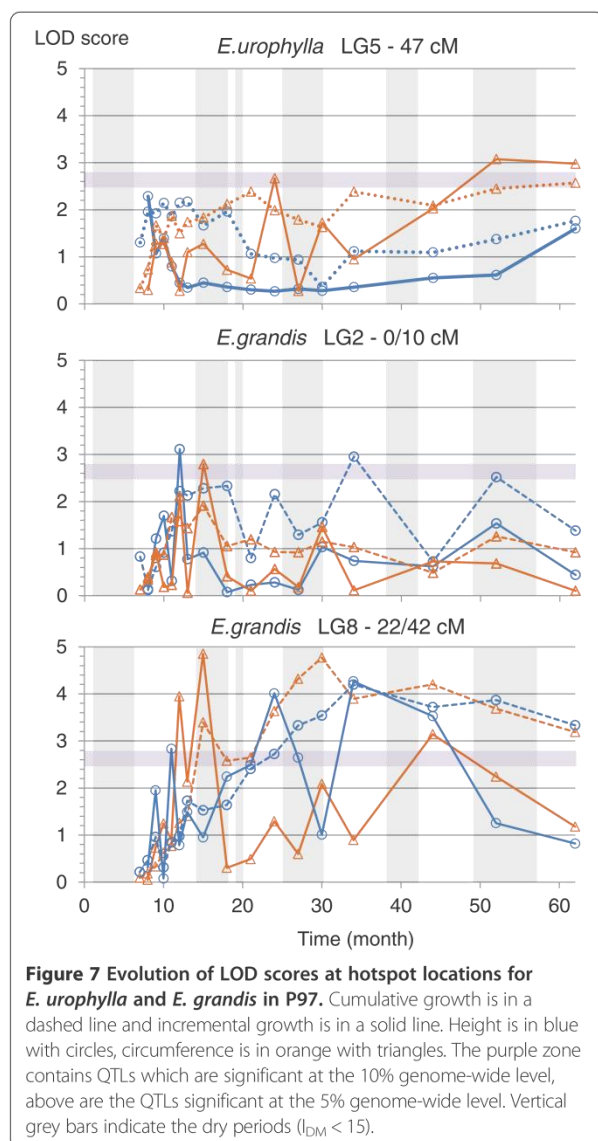
Trials P93 and P97 were compared using the growth traits (cumulative and incremental growth), GCPs, as well as specific predicted values (Figure 8). Proportionally to the number of traits, fewer QTL were detected for *E. urophylla* in P97 compared to P93. The proportion of QTLs detected for late growth traits (after 24 months) was higher than for early growth traits for *E. grandis* in both trials and was equivalent for *E. urophylla* in both trials. This may reflect the higher heritability of late growth traits compared with early growth as shown for the same family with P98. Considering P93 and P97, QTLs were detected on all LGs except on LG9 and LG11. However, for LG1 and LG7, QTLs were detected in P93 only. QTLs for GCP were found on three maps (no QTL was detected for *E. grandis* in P93). They were mainly located in growth trait QTL-hotspots except for *E. urophylla* on LG3 and LG8 in P93 and LG10 in P97. QTLs for predicted values were all detected at the same location as other QTLs for growth. Thus, QTLs for predicted values described well the genetic architecture of growth for P97 allowing the comparison with P93.

Co-locating QTLs between P93 and P97 were found on LG2 and 5 for the *E. grandis* parent only, involving strong overlapping of confidence intervals. These two co-localizations involved hotspots of one of two trials but represented few of all the detected QTLs. Indeed a Jaccard Index ( $I_j$ , [86]) was calculated between the two trials (P93 and P97) for *E. grandis* ( $I_j = 12.5\%$ ). This difference in the genetic architecture of growth traits



**Figure 6 QTL results for each parental tree of the P97 trial.** The number of QTLs in different genomic regions and their percentage of explained variance (PEV, %) range are given for cumulative growth and growth increment. A colour code indicates the % of QTLs detected in the whole map for each type of trait (from grey = no QTL to dark red = LG with the maximum number of QTLs).





between P93 and P97 suggested a strong QTL x environment interaction (QTL x E) that was tested using the predicted values obtained for P97 based on growth curve modelling. Thus, using closest markers already associated with QTLs, a two way analysis of variance was conducted to test QTL x E interaction for growth traits at ages 14, 26, 39, 51 and 59 months for P93 and predicted values at the same ages for P97. GCP were also used for both trials (Additional file 11). We found that most of tested markers revealed a significant interaction with the environment (trial) for both parents, *E. urophylla* 60% and 80%, *E. grandis* 40% and 90% for P93 and P97 respectively. For both trials, QTL x E was more significant for *E. grandis* than for *E. urophylla*.

## Discussion

### Incremental growth at the early stage is related to water availability and determines biomass production at rotation age

We combined the measurement of growth on a sub-annual scale with a monthly characterisation of climate variation using an aridity index ( $I_{DM}$ ), and first showed that seasonal changes in water availability affected secondary and primary growth at the early stage of tree development. Indeed,  $I_{DM}$  was found to be strongly correlated with radial growth increment before two years old, while it was less related to growth after two years. Besides, the correlation at early stage was more significant for radial growth ( $r = 0.69$ ,  $p < 0.01$ ) than for vertical growth ( $r = 0.5$ ,  $p < 0.01$ ), indicating that processes underlying secondary growth are more prone to environmental variation than those involved in primary growth. Other studies have already reported a relationship between apical or radial incremental growth and rainfall on different time scales (day, month and season) in different species [87-89]. In this study, we showed that the first two years corresponded to the highest biomass productivity period, during which 60% of the final height and 65% of final circumference was reached, representing a biomass productivity of  $20-27 \text{ m}^3 \cdot \text{ha}^{-1} \cdot \text{year}^{-1}$ . Therefore, the early growing period constitutes a key period with a strong impact on final growth, as illustrated by the high correlations values between each measurement point and the final growth at 60 months old. In tropical eucalypt plantations, it was found that the maximum growth rate occurred around two years, followed by a decline in production after canopy closure [90-92]. Therefore, our results suggest that climatic conditions underlying the aridity index (mainly rainfall) have a strong impact on early growth response and consequently determine biomass production at rotation age.

### Early growth response to water availability is genetically variable

The small sample size (60 genotypes) used to calculate  $H^2$  limited the accuracy of the estimation. However, the comparison of time trends in  $H^2$  between cumulative and incremental growths highlighted contrasting patterns over time which was significant only for Ht. Indeed, cumulative growth displayed an asymptotic increase up to rotation age, contrary to incremental growth which showed high temporal heterogeneity. For *Eucalyptus* grown in Congo,  $H^2$  for cumulative growth calculated using a wider genetic background were found to range from 0.10 to 0.5 [27] and to stabilize after two or three years. The same trend in  $H^2$  for cumulative growth was found in *Eucalyptus* in other environments [93-95]. Interestingly, this stabilization of cumulative growth heritabilities occurred at the same time as the decline in wood biomass productivity [92], with a

Species	Trials	Trait type	Nb trait	Nb QTL	Nb zone	PEV	LG										
							1	2	3	4	5	6	7	8	9	10	11
<i>E.uro</i>	P93	Growth	11	12	5	5.5 - 11.5											
		GCP	4	7	5	7.2 - 11.2											
	P97	Growth	21	25	7	5 - 8.9											
		GCP	1	1	1	6.9											
		PV	7	7	3	5.5 - 7.7											
<i>E.gra</i>	P93	Growth	8	9	6	4.2 - 11.3											
		GCP	-	-	-	-											
	P97	Growth	42	48	7	5.1 - 13											
		GCP	3	4	2	4.1 - 9.5											
		PV	16	19	2	5 - 13.7											

**Figure 8 QTL results for each parental tree in the two trials, P93 and P97.** The number of QTLs in different genomic regions and their percentage of explained variance (PEV, %) range are provided for growth traits (cumulative growth and growth increment), growth curve parameters (GCP) and predicted values (PV) from the monomolecular model. A colour code indicates the % of QTLs detected in the whole map for each type of trait (from grey = no QTL to dark red = LG with the maximum number of QTLs).

decrease in or stabilization of the  $G \times E$  interaction [95-97]. It could thus be suggested that the integrative nature of cumulative growth at a mature stage aggregates past genetic responses to environmental variations.

The heterogeneity of  $H^2$  for incremental growth observed for both height and radial growth suggests a variation in the impact of environmental conditions on growth response. The slowest values of incremental growth heritability were observed during the third dry period. In *Pinus Pinaster*, such heterogeneous trends in  $H^2$  have been shown for yearly incremental growths [98] and the variability described was significantly linked to environmental effects (year effects). On a lower scale, Parisi [99] also reported a heterogeneous pattern for narrow and broad-sense heritabilities for incremental height growth in *Pinus taeda*. In addition to the contrast between cumulative and incremental growth,  $H^2$  patterns for growth increment revealed opposite patterns during the 25th and 35th months between primary and secondary growth responses. Costa and Durel [98] also reported a clear difference between primary and secondary growth increments and suggested more stable genetic control for secondary growth response.

Between-tree variation in this early growth response, depicted by the distribution of phenotypic correlations between growth increments and  $I_{DM}$  during early growth, suggested a genetic effect as a major component in this response. Other studies have already suggested a genetic effect in response to water availability. Drew *et al.* [17] observed such contrasting behaviour in response to climatic variations (drought) between different *Eucalyptus* hybrids i.e. an *E. grandis*  $\times$  *E. urophylla* susceptible to drought with intermittent growth, and a less susceptible *E. grandis*  $\times$  *E. camaldulensis* maintaining a continuous growth. Similar results were found between *E. alba* and

*E. urophylla*  $\times$  *E. grandis* in Congo [18], associated with a significant genotype  $\times$  environment interaction for traits related to height growth and carbon isotope discrimination ( $\delta^{13}C$ ). These genetic differences in environmental susceptibility were also suggested by Bouvet *et al.* [100] comparing different eucalyptus clones. Interestingly, in the present study, we found a QTL for the correlation between  $I_{DM}$  and radial growth increment, suggesting that early growth response to water availability is indeed genetically controlled.

#### The genetic architecture of final growth is driven by the early growth response to the environment

Given the high macro-synteny and colinearity between eucalyptus genomes, we were able to compare the position of the mapped QTLs to that detected in other studies in different genotypes of the same species or different species. This comparison enabled us to identify two LGs where growth QTLs were frequently detected [39,41-43,52]. As already mentioned by Thumma *et al.* [43], and supported by our result, LG5 was found to contain QTLs for height and radial growth in different *Eucalyptus* species [39,41-43,52] as well as QTLs for foliar chemicals [101]. To a lesser extent, LG8, which was associated with both incremental and cumulative growth in our study, was also associated with growth-QTLs in three other studies [41,43,52], as well as QTLs for rooting ability [102] and QTLs for resistance to biotic stresses [103]. In addition to these two main LGs, and in agreement with the oligogenic model, i.e. few genes with large effects detected as QTLs and many genes with small effects controlled quantitative traits [104], QTLs were found to be distributed across the other nine chromosomes [39,41-44,52].

In this study, a similarity index revealed that most of the growth QTLs were not shared between two



experimental trials showing different environmental trajectories (see Figure 1). The statistical power of QTL detection and of estimated effects is directly related to sample size, so called Beavis effect [105]. Therefore QTLs with intermediate or small effects are likely to be not detected with a small sample size (e.g. 100 genotypes). Thus one cannot rule out the fact that QTL instability between P93 and P97 is the result of a rather low sample size: 200 and 190 genotypes for P93 and P97, respectively. This lack of QTL stability could also reveal true biological effect, i.e. a genotypic sensitivity to environmental conditions, as supported by the significant QTL  $\times$  E interaction found in both parents, and more particularly in *E. grandis*. Interestingly, the two studied species display contrasting adaptability to Congolese conditions with *E. urophylla* better adapted than *E. grandis* [4]. The instability of growth-QTLs between trials also contrasted with the QTL stability found for other traits measured on the same offspring (wood properties and  $\delta^{13}C$ , data not shown). This difference in sensitivity to environmental variations is generally used to explain the greater PEV and larger number of wood property QTLs in comparison to growth related traits [39,42,106]. It has been found that growth-QTLs often presented a higher level of interaction with the environment [47,48,107] than wood property-QTLs [41,52].

The temporal analysis of QTL detection showed the presence of early stage, intermediate and late QTLs. Among these three patterns of QTL expression, a large proportion of QTLs associated with cumulative growth exhibited early or intermediate patterns: 50% (P93) and 75% (P97) for *E. urophylla* and 75% (P93) and 33% (P97) for *E. grandis*. These three QTL patterns were also reported by Emebiri *et al.* [108] on three-year-old pines. This instability over time is frequently reported in forest tree literature in relation to low heritabilities associated with growth. In *E. globulus*, a study also reported a lack of QTL stability over time for diameter at breast height between two and six years [42]. In the same species, Bundock *et al.* [41] found partial QTL stability for stem diameter with the same location of LOD peaks between the diameter at two and six years old. Along the same lines, Thumma *et al.* [43] also reported relatively low stability for growth QTLs (height and diameter) over time (at two different ages) in one family. In our study, partial stability was found on LG2 (*E. grandis*) and LG5 (*E. urophylla*). Moreover, stable QTLs after two or three years old were found on LG6 (*E. urophylla*), LG8 (*E. grandis*) and LG10 (*E. grandis*). QTL stability over the years has also been reported in gymnosperms [50,109] but their measurements were limited to the young stage (two- to ten-year-old trees) and did not therefore reflect the genetic architecture of the whole growth trajectory.

To our knowledge, our study provides the most comprehensive view of the global instability of growth genetic architecture in a forest tree species thanks to a detailed analysis of the temporal and environmental dynamics of growth over the whole rotation period. Indeed, the fine scale characterization of both growth (multiple measurements) and the environment (aridity index) enabled a precise QTL dissection from early to late growth. In the literature, growth increments are mostly calculated on a yearly basis. Consequently, differences in genetic determinism between cumulative and incremental are less noticeable [109,110]. In our study, the wider variety of growth responses captured by short incremental data was instrumental to characterizing the genetic architecture of growth and depicting its temporal plasticity. More remarkably, thanks to the fine characterization of growth, we found that all QTLs detected at the mature stage for cumulative growth, co-localized with QTLs for early growth increment on both parents (LG5 for *E. urophylla* and LG2, LG8 for *E. grandis*). This result, which goes far beyond a simple age-age correlation analysis, suggests a partly common genetic determinism between early growth response to environmental variation and final growth. The QTL for the correlation with  $I_{DM}$  detected on LG2 in *E. grandis*, as well as the pattern of LOD scores for QTLs detected for mature growth on P93 (Additional file 12), corroborate this assumption. To sum up, our study shows that the variability of growth plasticity observed at the young stage partly determined the pattern of cumulative growth genetic architecture at rotation age.

## Conclusion

This study showed differences at phenotypic and genetic levels (phenotypic correlations, broad sense heritability, QTLs) between cumulative and incremental growth, for both primary and secondary growth in a fast growing tree species. Growth increments (especially radial growth) were found to be highly responsive to climatic variations (seasonal changes in water availability) over the two first years, contrary to cumulative growth. However, QTL analysis showed a common genetic basis between early growth increments and final growth. Moreover, strong instability was found in growth-QTLs between the two trials, which was also supported by the analysis of QTL  $\times$  E. These results suggest that early growth responses (apical and radial growth) to soil water availability partially shape the genetic determinism of late growth and thus growth trajectories. Specific genomic regions were also found for growth increments and growth curve parameters, showing the merits of a combined approach with sequential (increment) and global

(curve parameters) growth traits to dissect its genetic determinism. Our results highlight the importance of considering growth as a compound trait for a better understanding of its genetic bases, and suggest to combine selection on early and cumulative growth to improve the ability of trees to produce wood biomass in an optimal way.

## Additional files

**Additional file 1: Growth trajectories of four individuals (1-4) randomly selected from the full-sib progenies of each trial.** The circumference growth trajectories were fitted by the monomolecular model (purple curve). The mean growth curve for all individuals is represented with a blue line.

**Additional file 2: Mean trajectory for cumulative height and circumference for the three trials, P93, P97 and P98.** The thick coloured bars along the x-axis represent dry periods ( $IDM < 15$ ) for each trial (red for P93, blue for P97 and green for P98).

**Additional file 3: Summary of phenotypic data for each trial, P93, P97 and P98.** Traits, number of values, minimum (Min), maximum (Max), mean, standard deviation (SD) and coefficient of phenotypic variation (CVp).

**Additional file 4: Evolution over time of the CVp for cumulative and incremental growth for P97 and P98.** The thick grey bars along the x-axis represent dry periods ( $IDM < 15$ ). Cumulative growth is in dashed lines and incremental growth is in solid lines. Height is in blue with circles, circumference is in orange with triangles.

**Additional file 5: Matrix of Pearson correlation coefficients.** For all traits: cumulative growth, growth increment, growth curve parameters and predicted values for P93 (A), P97 (B) and P93 (C). Only significant correlation values ( $p$ -value  $< 0.05$ ) are displayed.

**Additional file 6: Phenotypic correlations between final growth and current growth for the three trials.** Cumulative growth is in dashed lines and incremental growth is in solid lines. Height is in blue with circles, circumference is in orange with triangles. Vertical grey bars indicate the dry periods.

**Additional file 7: Broad sense heritabilities ( $H^2$ ) with confidence interval (CI) for growth related traits.**

**Additional file 8: Genetic maps of *E. urophylla* and *E. grandis* for P93 and P97.** The LG nomenclature corresponds to the one defined by Brondani et al. [81]. Distances between markers are indicated in centiMorgans (cM) by a scale on the left side. Common markers between parental maps for P93 and P97 are in bold. Orthologous regions between the P93 and P97 maps are shown in red (*E. urophylla*) or orange (*E. grandis*) hatching. Common markers between the two parents of P93 are in bold and grey, orthologous regions are represented by a grey bar segment. Accessory markers are positioned near the closest framework marker with the distance in brackets. QTLs are represented by boxes extended by lines representing 95% Bayes credible intervals. If multiple QTLs of the same trait type for different ages co-localize, only one QTL is represented with the largest interval and (\*) indicates if one or more QTLs were significant at 5% whole-genome level. Accessory markers with an (\*) were distorted ( $p < 0.01$ ). Traits have been explained in the text.

**Additional file 9: QTLs detected in P93 by composite interval mapping in *E. urophylla* and *E. grandis*.** The trait, the number of genotype used for analysis (n), the linkage group (LG), the position, the LOD value (\* = significant at 5% genome-wide level), the percentage of variance explained by the QTL (PEV, %) and the difference between the two QTL allele effects (D) are indicated.

**Additional file 10: QTLs detected in P97 by composite interval mapping in *E. urophylla* and *E. grandis*.** The trait, the number of genotype used for analysis (n), the linkage group (LG), the position, the LOD value (\* = significant at 5% genome-wide level), the percentage of variance explained by the QTL (PEV, %) and the difference between the two QTL allele effects (D) are indicated.

**Additional file 11: Analysis of variance for markers associated with a QTL in P93 (A) or P97 (B).** Similar measurement ages or predicted values were used to analyse the two trials. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

**Additional file 12: Evolution of LOD scores at hotspot locations on *E. urophylla* and *E. grandis* for P93.** Height is in blue with a circle, circumference is in orange with a triangle. The purple zone contains QTLs which are significant at the 10% genome-wide level, above are the QTLs significant at the 5% genome wide level. Vertical grey bars indicate the dry periods ( $IDM < 15$ ).

## Competing interests

The authors have no competing interests to declare.

## Authors' contributions

FS: DNA extraction and genotyping. JMB, PV: Field trial management. JB: growth modelling, linkage analysis, QTL mapping. JB, JMG, CP, JMB: manuscript preparation. JMG: Funding and overall supervision. All authors read and approved the final manuscript.

## Acknowledgements

The authors would like to thank the technicians at CRDPI (Centre de Recherches sur la Durabilité des Plantations Industrielles located in the Republic of Congo) for trial establishment and follow-up. The authors acknowledge two anonymous reviewers for their constructive comments and suggestions on the manuscript. This study was supported by grants from FEDER (Fonds Européen de Développement Régional, ABIOPEN project, n° Presage 32973). JB received a PhD fellowship from the BIOS department of CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement).

## Author details

<sup>1</sup>CIRAD, UMR AGAP, F-33612 Cestas, France. <sup>2</sup>INRA, UMR BIOGECO, F-33612 Cestas, France. <sup>3</sup>CIRAD, UMR AGAP, F-34398 Montpellier, France. <sup>4</sup>CIRAD, UMR AGAP, Pointe Noire, Congo. <sup>5</sup>CRDPI, BP 1291 Pointe Noire, Rep. of Congo.

Received: 29 April 2013 Accepted: 14 August 2013

Published: 26 August 2013

## References

1. FAO: *State of the world's forests, 2009*. Rome: Food and Agriculture Organization of the United Nations; 2009.
2. FAO: *State of the world's forests: 2007*. Rome: Food and Agriculture Organization of the United Nations; 2007.
3. Coppen JJW: *Eucalyptus: The Genus Eucalyptus*. London: Taylor & Francis; 2002.
4. Vigneron P: *Création et amélioration de variétés hybrides d'Eucalyptus au Congo. Bois et forêts des tropiques* 1992, **234**:29–42.
5. Denison NP, Kietzka JE: *The use and importance of hybrid intensive forestry in South Africa. South African Forestry Journal* 1993, **165**:55–60.
6. Kozowski TT, Pallardy SG: *Growth control in woody plants*. New York: Academic Press; 1997.
7. Barton MK: *Twenty years on: the inner workings of the shoot apical meristem, a developmental dynamo. Dev Biol* 2010, **341**(1):95–113.
8. Lachaud S, Cateson A-M, Bonnemaison J-L: *Structure and functions of the vascular cambium. Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie* 1999, **322**(8):633–650.
9. Vanclay JK: *Growth models for tropical forests: a synthesis of models and methods. Forest Sci* 1995, **41**(1):7–42.
10. Vanclay JK, Skovsgaard JP: *Evaluating forest growth models. Ecol Model* 1997, **98**(1):1–12.
11. Zeide B: *The science of forestry. J Sustain For* 2008, **27**(4):345–473.
12. Landsberg JJ, Sands P: *Physiological ecology of forest production: principles, processes and models*. London: Elsevier Science; 2010.
13. Mäkelä A, Landsberg J, Ek AR, Burk TE, Ter-Mikaelian M, Ågren GI, Oliver CD, Puttonen P: *Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiol* 2000, **20**(5–6):289–298.
14. Landsberg J: *Physiology in forest models: history and the future. FBMIS* 2003, **1**:49–63.



15. Almeida AC, Landsberg JJ, Sands PJ, Ambrogi MS, Fonseca S, Barddal SM, Bertolucci FL: **Needs and opportunities for using a process-based productivity model as a practical tool in Eucalyptus plantations.** *For Ecol Manage* 2004, **193**(1-2):167-177.
16. Fontes L, Landsberg J, Tome J, Tome M, Pacheco CA, Soares P, Araujo C: **Calibration and testing of a generalized process-based model for use in Portuguese eucalyptus plantations.** *Can J For Res-Rev Can Rech For* 2006, **36**(12):3209-3221.
17. Drew D, Downes G, Grzeskowiak V, Naidoo T: **Differences in daily stem size variation and growth in two hybrid eucalypt clones.** *Trees - Structure and Function* 2009, **23**(3):585-595.
18. Villar E: *Plasticité phénotypique et moléculaire de deux clones d'eucalyptus sous contrainte hydrique au champ*, Ph.D. thesis. Bordeaux: Université de Bordeaux I; 2011.
19. Namkoong G, Kang H: **Quantitative Genetics of Forest Trees.** In *Plant Breeding Reviews*. New Jersey: John Wiley & Sons, Inc; 1990:139-188.
20. Cornelius J: **Heritabilities and additive genetic coefficients of variation in forest trees.** *Can J Forest Res* 1994, **24**(2):372-379.
21. Hamilton M, Potts B: **Review of Eucalyptus nitens genetic parameters.** *N Z J For Sci* 2008, **38**(1):102-119.
22. Volker PW, Potts BM, Borralho NMG: **Genetic parameters of intra- and inter-specific hybrids of Eucalyptus globulus and E-nitens.** *Tree Genetics & Genomes* 2008, **4**(3):445-460.
23. Bouvet JM, Saya A, Vigneron P: **Trends in additive, dominance and environmental effects with age for growth traits in Eucalyptus hybrid populations.** *Euphytica* 2009, **165**(1):35-54.
24. Retief ECL, Stanger TK: **Genetic parameters of pure and hybrid populations of Eucalyptus grandis and E-uophylla and implications for hybrid breeding strategy.** *South Forests* 2009, **71**(2):133-140.
25. Franklin EC: **Model relating levels of genetic variance to stand development of 4 north-american conifers.** *Silvae genetica* 1979, **28**(5-6):207-212.
26. Cotterill PP, Dean CA: **Changes in the genetic-control of growth of radiata pine to 16 years and efficiencies of early selection.** *Silvae genetica* 1988, **37**(3-4):138-146.
27. Bouvet JM, Vigneron P: **Age trends in variances and heritabilities in Eucalyptus factorial mating designs.** *Silvae genetica* 1995, **44**(4):206-216.
28. Stackpole D, Vaillancourt R, De Aguiar M, Potts B: **Age trends in genetic parameters for growth and wood density in Eucalyptus globulus.** *Tree Genetics & Genomes* 2010, **6**(2):179-193.
29. Xiang B, Li B, Isik F: **Time trend of genetic parameters in growth traits of Pinus taeda L.** *Silvae genetica* 2003, **52**(3-4):114-121.
30. Kremer A: **Predictions of age-age correlations of total height based on serial correlations between height increments in Maritime pine (Pinus pinaster Ait.).** *TAG Theor Appl Genet* 1992, **85**(2):152-158.
31. Gwaze DP, Bridgwater FE, Byram TD, Woolliams JA, Williams CG: **Predicting age-age genetic correlations in tree-breeding programs: a case study of Pinus taeda L.** *Theor Appl Genet* 2000, **100**(2):199-206.
32. Martinez Meier A, Sanchez G, Dalla L, Salda G, Pastorino MJM, Gautry JY, Gallo LA, Rozenberg P: **Genetic control of the tree-ring response of Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco) to the 2003 drought and heat-wave in France.** *Ann For Sci* 2008, **65**(1):102.
33. Marron N, Ricciotti L, Bastien C, Beritognolo I, Gaudet M, Paolucci I, Fabbri F, Salani F, Dillen SY, Ceulemans R, et al: **Plasticity of growth and biomass production of an intraspecific Populus alba family grown at three sites across Europe during three growing seasons.** *Can J For Res-Rev Can Rech For* 2010, **40**(10):1887-1903.
34. Costae Silva J, Potts BM, Dutkowski GW: **Genotype by environment interaction for growth of Eucalyptus globulus in Australia.** *Tree Genetics & Genomes* 2006, **2**(2):61-75.
35. Dutkowski GW, Potts BM: **Genetic variation in the susceptibility of Eucalyptus globulus to drought damage.** *Tree Genetics & Genomes* 2012, **8**(4):757-773.
36. Ritland K, Krutovsky KV, Tsumura Y, Pelgas B, Isabel N, Bousquet J: **Genetic mapping in conifers.** In *Genetics, genomics and breeding of conifers*. Edited by Plomion C, Bousquet J, Kole C. New York: Science Publishers & CRC Press; 2011:1121-1134.
37. Rae AM, Street NR, Rodriguez-Acosta M: **Populus trees.** In *Forest Trees*. UK: Springer; 2007:1-28.
38. Myburg AA, Potts BM, Marques CM, Kirst M, Gion JM, Grattapaglia D, Grima-Pettenatti J: *Eucalypts*. Heidelberg, Germany: Springer-Verlag GmbH; 2007.
39. Grattapaglia D, Bertolucci F, Penchel R, Sederoff R: **Genetic mapping of quantitative trait loci controlling growth and wood quality traits in Eucalyptus grandis using a maternal half-sib family and RAPD markers.** *Genetics* 1996, **144**:1205-1214.
40. Verhaegen D, Plomion C, Gion JM, Poitel M, Costa P, Kremer A: **Quantitative trait dissection analysis in Eucalyptus using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages.** *TAG Theor Appl Genet* 1997, **95**(4):597-608.
41. Bundock PC, Potts BM, Vaillancourt RE: **Detection and stability of quantitative trait loci (QTL) in Eucalyptus globulus.** *Tree Genetics & Genomes* 2008, **4**(1):85-95.
42. Freeman JS, Whitlock SP, Potts BM, Vaillancourt RE: **QTL influencing growth and wood properties in Eucalyptus globulus.** *Tree Genetics & Genomes* 2009, **5**(4):713-722.
43. Thumma BR, Baltunis BS, Bell JC, Emebiri LC, Moran GF, Southerton SG: **Quantitative trait locus (QTL) analysis of growth and vegetative propagation traits in Eucalyptus nitens full-sib families.** *Tree Genetics & Genomes* 2010, **6**(6):877-889.
44. Kullán AR, Van Dyk M, Hefer C, Jones N, Kanzler A, Myburg A: **Genetic dissection of growth, wood basic density and gene expression in interspecific backcrosses of Eucalyptus grandis and E. urophylla.** *BMC Genet* 2012, **13**(1):60.
45. Ma C-X, Casella G, Wu R: **Functional Mapping of Quantitative Trait Loci Underlying the Character Process: A Theoretical Framework.** *Genetics* 2002, **161**(4):1751-1762.
46. Wu R, Lin M: **Functional mapping: how to map and study the genetic architecture of dynamic complex traits.** *Nat Rev Genet* 2006, **7**(3):229-237.
47. Rae A, Pinel M, Bastien C, Sabatti M, Street N, Tucker J, Dixon C, Marron N, Dillen S, Taylor G: **QTL for yield in bioenergy Populus: identifying G x E interactions from growth at three contrasting sites.** *Tree Genetics & Genomes* 2008, **4**(1):97-112.
48. Dillen SY, Storme V, Marron N, Bastien C, Neyrinck S, Steenackers M, Ceulemans R, Boerjan W: **Genomic regions involved in productivity of two interspecific poplar families in Europe. 1. Stem height, circumference and volume.** *Tree Genetics & Genomes* 2009, **5**(1):147-164.
49. Rönnerberg-Wästljung AC, Glynn C, Weih M: **QTL analyses of drought tolerance and growth for a Salix dasyclados x Salix viminalis hybrid in contrasting water regimes.** *TAG Theor Appl Genet* 2005, **110**(3):537-549.
50. Pelgas B, Bousquet J, Meimans P, Ritland K, Isabel N: **QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments.** *BMC Genomics* 2011, **12**(1):145.
51. Freeman J, Potts B, Downes G, Thavamanikumar S, Pilbeam D, Hudson C, Vaillancourt R: **QTL analysis for growth and wood properties across multiple pedigrees and sites in Eucalyptus globulus.** *BMC Proc* 2011, **5**(Suppl 7):O8.
52. Freeman JS, Potts BM, Downes GM, Pilbeam D, Thavamanikumar S, Vaillancourt RE: **Stability of quantitative trait loci for growth and wood properties across multiple pedigrees and environments in Eucalyptus globulus.** *New Phytol* 2013, **198**(4):1121-1134.
53. Teixeira J, Missiaggia A, Dias D, Scarpinati E, Viana J, Paula N, Paula R, Bonine C: **QTL analyses of drought tolerance in Eucalyptus under two contrasting water regimes.** *BMC Proc* 2011, **5**(Suppl 7):P40.
54. Grattapaglia D, Plomion C, Kirst M, Sederoff RR: **Genomics of growth traits in forest trees.** *Curr Opin Plant Biol* 2009, **12**(2):148-156.
55. Gion J-M, Rech P, Grima-Pettenatti J, Verhaegen D, Plomion C: **Mapping candidate genes in Eucalyptus with emphasis on lignification genes.** *Mol Breed* 2000, **6**(5):441-449.
56. Gion J-M, Carouche A, Deweer S, Bedon F, Pichavant F, Charpentier J-P, Bailleres H, Rozenberg P, Carocha V, Ognouabi N, et al: **Comprehensive genetic dissection of wood properties in a widely-grown tropical tree: eucalyptus.** *BMC Genomics* 2011, **12**(1):301.
57. Doyle J, Doyle J: **Isolation of plant DNA from fresh tissue.** *Focus* 1990, **12**:13-15.
58. Mareschal L, Nzila JDD, Turpault MP, Thongo M'Bou A, Mazombou JC, Bouillet JP, Ranger J, Laclau JP: **Mineralogical and physico-chemical properties of Ferralic Arenosols derived from unconsolidated Plio-Pleistocene deposits in the coastal plains of Congo.** *Geoderma* 2011, **162**(1-2):159-170.
59. Kottek M, Grieser J, Beck C, Rudolf B, Rubel F: **World Map of the Koppen-Geiger climate classification updated.** *Meteorol Z* 2006, **15**(3):259-263.
60. De Martonne E: *Une Nouvelle fonction climatologique. L'indice d'aridité*. Paris: Impr. Gauthier-Villars; 1926.



61. Von Bertalanffy L: Quantitative Laws in Metabolism and Growth. *Q Rev Biol* 1957, **32**(3):217–231.
62. Richards FJ: A Flexible Growth Function for Empirical Use. *J Exp Bot* 1959, **10**(2):290–301.
63. Pienaar LV, Turnbull KJ: The Chapman-Richards Generalization of Von Bertalanffy's Growth Model for Basal Area Growth and Yield in Even - Aged Stands. *Forest Sci* 1973, **19**(1):2–22.
64. Zeide B: Analysis of Growth Equations. *Forest Sci* 1993, **39**(3):594–616.
65. Vanclay JK: Modelling forest growth and yield: applications to mixed tropical forests. Wallingford UK: CAB International; 1994.
66. Fekedulegn D, Mac Siurtain MP, Colbert JJ: Parameter estimation of nonlinear growth models in forestry. *Silva Fennica* 1999, **33**(4):327–336.
67. Lei YC, Zhang SY: Features and partial derivatives of Bertalanffy-Richards growth model in forestry. *Nonlinear Analysis: Modelling and Control* 2004, **9**(1):65–73.
68. Bouvet JM: Evolution de la variabilité avec l'âge et corrélation juvénile-adulte dans des populations d'eucalyptus. Ph.D. Paris: Institut National Agronomique; 1995.
69. Amaro A, Reed D, Tome M, Themido I: Modeling dominant height growth: Eucalyptus plantations in Portugal. *Forest Sci* 1998, **44**(1):37–46.
70. Saint-Andre L, Laclau J-P, Deleporte P, Ranger J, Gouma R, Saya A, Joffre R: A Generic Model to Describe the Dynamics of Nutrient Concentrations within Stemwood across an Age Series of a Eucalyptus Hybrid. *Ann Bot* 2002, **90**(1):65–76.
71. Pinheiro JC, Bates DM: Mixed-Effects Models in S and S-Plus. New York: Springer; 2009.
72. Pinheiro J, Bates D, Debroy S, Sarkar D, the R Development Core Team: nlme: Linear and Nonlinear Mixed Effects Models. In *R package version 3.1-110 edn*; 2013.
73. Hadfield JD: MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *J Stat Softw* 2010, **33**(2):1–22.
74. Nyquist WE, Baker RJ: Estimation of heritability and prediction of selection response in plant populations. *Crit Rev Plant Sci* 1991, **10**(3):235–322.
75. Grattapaglia D, Sederoff R: Genetic Linkage Maps of Eucalyptus grandis and Eucalyptus urophylla Using a Pseudo-Testcross Mapping Strategy and RAPD Markers. *Genetics* 1994, **137**(4):1121–1137.
76. Stam P: Construction of integrated genetic linkage maps by means of a new computer package: join map. *Plant J* 1993, **3**(5):739–744.
77. Van Ooijen JW: JoinMap 4.1, Software for the Calculation of Genetic Linkage Maps in Experimental Populations. Wageningen, Netherlands: Kyazma BV; 2011.
78. Jansen J, De Jong AG, Van Ooijen JW: Constructing dense genetic linkage maps. *TAG Theor Appl Genet* 2001, **102**(6):1113–1122.
79. Van Ooijen JW: Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. *Genet Res* 2011, **93**(5):343–349.
80. Kosambi DD: The estimation of map distances from recombination values. *Ann Hum Genet* 1943, **12**(1):172–175.
81. Brondani R, Williams E, Brondani C, Grattapaglia D: A microsatellite-based consensus linkage map for species of Eucalyptus and a novel set of 230 microsatellite markers for the genus. *BMC Plant Biol* 2006, **6**(1):20.
82. Broman KW, Wu H, Sen S, Churchill GA: R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 2003, **19**(7):889–890.
83. Sen S, Churchill GA: A Statistical Framework for Quantitative Trait Mapping. *Genetics* 2001, **159**(1):371–387.
84. Churchill GA, Doerge RW: Empirical Threshold Values for Quantitative Trait Mapping. *Genetics* 1994, **138**(3):963–971.
85. Voorrips RE: MapChart: software for the Graphical Presentation of Linkage Maps and QTLs. *J Hered* 2002, **93**(1):77–78.
86. Jaccard P: Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. *Bulletin de la Société Vaudoise des Sciences Naturelles* 1901, **37**:241–272.
87. Zweifel R, Zimmermann L, Zeugin F, Newbery DM: Intra-annual radial growth and water relations of trees: implications towards a growth mechanism. *J Exp Bot* 2006, **57**(6):1445–1459.
88. Craven D, Dent D, Braden D, Ashton MS, Berlyn GP, Hall JS: Seasonal variability of photosynthetic characteristics influences growth of eight tropical tree species at two sites with contrasting precipitation in Panama. *For Ecol Manage* 2011, **261**(10):1643–1653.
89. Gutierrez E, Campelo F, Camarero JJ, Ribas M, Muntan E, Nabais C, Freitas H: Climate controls act at different scales on the seasonal pattern of Quercus ilex L. stem radial increments in NE Spain. *Trees-Struct Funct* 2011, **25**(4):637–646.
90. Binkley D, Stape JL, Ryan MG, Barnard HR, Fownes J: Age-related decline in forest ecosystem growth: an individual-tree, stand-structure hypothesis. *Ecosystems* 2002, **5**(1):58–67.
91. Ryan MG, Binkley D, Fownes JH, Giardina CP, Senock RS: An experimental test of the causes of forest growth decline with stand age. *Ecol Monogr* 2004, **74**(3):393–414.
92. Ryan MG, Binkley D, Stape JL: Why don't our stands grow even faster? Control of production and carbon cycling in eucalypt plantations. *Southern Forests: a Journal of Forest Science* 2008, **71**(1):99–104(106).
93. Marques OG Jr, Andrade HB, Ramalho MAP: Assessment of the early selection efficiency in Eucalyptus cloeziana F. Muell. in the northwest of Minas Gerais state (Brazil). *Silvae genetica* 1996, **45**(5/6):359–361.
94. Wei X, Borralho NMG: Genetic control of growth traits of Eucalyptus urophylla S.T. Blake in south east China. *Silvae genetica* 1998, **47**(2-3):158–165.
95. Osorio LF, White TL, Huber DA: Age-age and trait-trait correlations for Eucalyptus grandis Hill ex Maiden and their implications for optimal selection age and design of clonal trials. *Theor Appl Genet* 2003, **106**(4):735–743.
96. Osorio LF, White TL, Huber DA: Age trends of heritabilities and genotype-by-environment interactions for growth traits and wood density from clonal trials of Eucalyptus grandis Hill ex Maiden. *Silvae genetica* 2001, **50**(1):30–37.
97. Kien ND, Jansson G, Harwood C, Thinh HH: Genetic control of growth and form in Eucalyptus urophylla in Northern Vietnam. *J Trop For Sci* 2009, **21**(1):50–56.
98. Costa P, Durel CE: Time trends in genetic control over height and diameter in maritime pine. *Can J Forest Res* 1996, **26**(7):1209–1217.
99. Parisi LM: Shoot elongation patterns and genetic control of second-year height growth in Pinus taeda L. using clonally replicated trials. Gainesville: University of Florida; 2006.
100. Bouvet JM, Vigneron P, Saya A: Phenotypic plasticity of growth trajectory and ontogenic allometry in response to density for Eucalyptus hybrid clones and families. *Ann Bot* 2005, **96**(5):811–821.
101. Henery ML, Moran GF, Wallis IR, Foley WJ: Identification of quantitative trait loci influencing foliar concentrations of terpenes and formylated phloroglucinol compounds in Eucalyptus nitens. *New Phytol* 2007, **176**(1):82–95.
102. Shepherd M, Kasem S, Lee D, Henry R: Mapping species differences for adventitious rooting in a Corymbia torelliana x Corymbia citriodora subspecies variegata hybrid. *Tree Genetics & Genomes* 2008, **4**(4):715–725.
103. Freeman JS, Potts BM, Vaillancourt RE: Few Mendelian genes underlie the quantitative response of a forest tree, Eucalyptus globulus, to a natural fungal epidemic. *Genetics* 2008, **178**(1):563–571.
104. Morton NE, MacLean CJ: Analysis of family resemblance. 3. Complex segregation of quantitative traits. *Am J Hum Gen* 1974, **26**(4):489–503.
105. Beavis WD: QTL analyses: power, precision, and accuracy. *Molecular dissection of complex traits* 1998, **1998**:145–162.
106. Markussen T, Fladung M, Achere V, Favre JM, Faivre-Rampant P, Aragones A, Perez DD, Harvengt L, Espinel S, Ritter E: Identification of QTLs controlling growth, chemical and physical wood property traits in Pinus pinaster. *Silvae genetica* 2003, **52**(1):8–15.
107. Wu R, Bradshaw HD, Stettler RF: Developmental quantitative genetics of growth in Populus. *Theor Appl Genet* 1998, **97**(7):1110–1119.
108. Emebiri LC, Devey ME, Matheson AC, Slee MU: Age-related changes in the expression of QTLs for growth in radiata pine seedlings. *TAG Theor Appl Genet* 1998, **97**(7):1053–1061.
109. Lerceteau E, Szmidt AE, Andersson B: Detection of quantitative trait loci in Pinus sylvestris L. across years. *Euphytica* 2001, **121**(2):117–122.
110. Emebiri LC: Detection and genetic mapping of quantitative trait loci influencing stem growth efficiency in radiata pine. Australia: Australian National University; 1997.

doi:10.1186/1471-2229-13-120

Cite this article as: Bartholomé et al.: Plasticity of primary and secondary growth dynamics in Eucalyptus hybrids: a quantitative genetics and QTL mapping perspective. *BMC Plant Biology* 2013 **13**:120.





## **4 High-resolution genetic linkage maps of *E. grandis* and *E. urophylla***



## **High-resolution genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* improve BRASUZ1 reference genome assembly and reveal segregation distortion regions.**

Jérôme BARTHOLOME <sup>1,2,3</sup>, Eric MANDROU <sup>2,3,4</sup>, André MABIALA <sup>5</sup>, Ibouniyamine NABIHOUDINE <sup>4</sup>, Christophe KLOPP <sup>4</sup>, Christophe PLOMION <sup>2,3</sup>, Jean-Marc GION <sup>1,2,3\*</sup>

<sup>1</sup> CIRAD, UMR AGAP, F-33612 Cestas, France

<sup>2</sup> INRA, UMR1202, BIOGECO, F-33610 Cestas, France

<sup>3</sup> Université de Bordeaux, UMR1202 BIOGECO, F-33170 Talence, France

<sup>4</sup> Plate-forme bio-informatique Genotoul, INRA, Biométrie et Intelligence Artificielle, BP 52627, 31326, Castanet-Tolosan Cedex, France

<sup>5</sup> CRDPI, BP 1291, Pointe Noire, Republic of Congo

\*Corresponding author: Jean-Marc GION



## Abstract

Genetic linkage maps are key tools in research as they provide the framework for many applications, such as quantitative trait loci analysis, and they support the assembly of genome sequences. Resequencing of the two parental genotypes of an interspecific cross between *Eucalyptus urophylla* and *E. grandis* was used to design an SNP array of 6,000 markers evenly distributed along the *Eucalyptus* BRASUZ1 reference genome sequence. The genotyping of 1,025 offspring enabled the construction of two high-resolution saturated genetic linkage maps of 822 and 886 cM, with a marker density of one marker every 0.45 and 0.5 cM for *Eucalyptus grandis* and *E. urophylla*, respectively. The comparison between linkage maps and the assembled genome sequence highlighted 85% of collinear regions, while 43 non-collinear regions and 13 non-synthetic regions were detected, mainly distributed in scaffold 1 and 5. The mapping data were also used to study the distribution of segregation distortion regions (SDR) and the underlying gene content for these two species. This report provides the most reliable linkage maps for *Eucalyptus*, an improved reference genome assembly and new insights into genomic regions associated with reproductive isolation.

**Keywords:** SNP array, Genetic mapping, Genome, Segregation distortion, *Eucalyptus*

## 4.1 Introduction

A century after the publication of the first genetic map of *Drosophila melanogaster* (Sturtevant, 1913), linkage maps have been established for many organisms (Botstein *et al.*, 1980; Koornneef *et al.*, 1983; Grattapaglia & Sederoff, 1994) and are still central in genetic research. Based on the recombination events between homologous chromosomes occurring during meiosis (i.e. crossover in prophase I, (Crismani *et al.*, 2013)), linkage maps provide a linear representation of order and distance between markers on chromosomes. Since the arrival of molecular markers (Schlotterer, 2004) and recent technological and bioinformatic advances facilitating their high-throughput discovery (Kumar *et al.*, 2012), genetic maps have enlarged their applications. Indeed, they constitute a powerful tool for characterizing the genome structure of non-sequenced species and play a major role in genome assembly and validation (e.g. scaffold anchoring) as illustrated in *Arabidopsis thaliana* (Kaul *et al.*, 2000), *Homo sapiens* (Lander *et al.*, 2001) and *Populus trichocarpa* (Tuskan *et al.*, 2006). Genetic



maps also provide insights into genome evolution through the analysis of synteny, collinearity and chromosomal rearrangements between species (Burt, 2002; Choi *et al.*, 2004; Krutovsky *et al.*, 2004; Hudson *et al.*, 2012b). Moreover, linkage maps form the basis for mapping and cloning quantitative trait loci (QTL) involved in the genetic control of traits of interest (Price, 2006; Salvi & Tuberosa, 2007; Mackay *et al.*, 2009; Wurschum, 2012).

The precision of recombination fraction estimates between linked markers, which results ultimately in map accuracy, is a crucial point for any of these applications. Several mapping algorithms have been developed to order the ever increasing number of markers and estimate genetic distances with more speed and reliability (Cheema & Dicks, 2009; Mollinari *et al.*, 2009). Merging methods have also been proposed to combine linkage maps (Stam, 1993; Peirce *et al.*, 2007; Endelman, 2011; Wu *et al.*, 2011; Ronin *et al.*, 2012) and build composite genetic maps. Although mapping algorithms have been shown to play an important role in mapping accuracy (Collard *et al.*, 2009; Mollinari *et al.*, 2009; Ronin *et al.*, 2010), it is primarily the number of recombination events captured in the mapping population, depending on the population type and its sample size, that determines mapping accuracy (Nelson, 2005; Ferreira *et al.*, 2006). In order to build high-resolution genetic linkage maps with a practical sample size, highly recombinant populations (e.g. IRILs (Ganal *et al.*, 2011) and sperm or pollen typing techniques (Yelina *et al.*, 2012) were developed. In outbred species, such as forest trees with a long generation time, mapping activities have mostly relied on existing full-sib families of rather small sample sizes (Kole, 2007) thereby limiting the accuracy and resolution of genetic maps. Such families were generated as part of breeding programmes for genetic parameter estimation. Although the number of markers is no longer a limiting factor (Sansaloni *et al.*, 2010; Geraldès *et al.*, 2013; Howe *et al.*, 2013) for building high-density linkage maps in forest trees, there is still room for progress in improving mapping accuracy with a larger population size, though at the expense of genotyping costs.

Genetic mapping approaches have been widely used to characterize the unsequenced genome of many forest tree species over recent decades (Kole, 2007; Neale & Kremer, 2011). For *Eucalyptus*, the most widely grown plantation hardwoods, numerous genetic maps have been built for different species (review in Table S4.1). Briefly, the first genetic maps were developed using an interspecific cross between *E. grandis* x *E. urophylla* with dominant markers following a double pseudo-testcross mapping strategy (Grattapaglia & Sederoff, 1994). Then, the development of simple sequence repeat markers (Brondani *et al.*, 1998;



Brondani *et al.*, 2006), as well as the identification of expressed sequenced tag polymorphisms (Gion *et al.*, 2000; Thamarus *et al.*, 2002), enabled broad-scale comparisons of genome organization across species (Gion *et al.*, in prep). More recently, a DArTs array was developed for the *Eucalyptus* and *Corymbia* species (Sansaloni *et al.*, 2010). Genetic maps with up to four thousand DArTs for *E. urophylla*, *E. grandis* and *E. globulus* (Hudson *et al.*, 2012a; 2012b; Kullán *et al.*, 2012; Petroli *et al.*, 2012) were constructed, specifying linkage group homologies as well as macro-synteny and collinearity between these species. However, most of the linkage maps were established with a rather small to moderate number of full-sib or back-cross offspring, ranging from 62 (Grattapaglia & Sederoff, 1994) to 503 (Hudson *et al.*, 2012b). Recently, a new era began for eucalyptus genomics with two main breakthroughs: (i) physical characterization of the *E. grandis* BRASUZ1 reference genome totaling 691 Mb distributed over 4,952 scaffolds. The main 11 scaffolds, representing the 11 chromosomes, totaled 88% (605.9 Mb) of the genome and 93% (33,917) of the predicted genes, and (ii) the development of a vast amount of polymorphism for different eucalyptus species (Novaes *et al.*, 2008; Sansaloni *et al.*, 2010; Grattapaglia *et al.*, 2011; Neves *et al.*, 2011) providing a considerable number of markers for high-resolution linkage mapping in this genus.

In this study, we developed genomic resources for the two parents of an interspecific cross between *E. urophylla* and *E. grandis* by resequencing their genome. The large amount of variants detected after the mapping paired-end reads on the BRASUZ1 genome enabled us to design a high-quality SNP genotyping array maximizing the coverage of the main 11 scaffolds. Genotyping of these markers on a large full-sib family totaling 1,025 F1s made it possible to achieve our three initial objectives: (i) construct the first high-resolution linkage maps of any tree species providing a framework for future map-based cloning activities of major-effect QTLs, (ii) test the effect of population size and number of markers on genetic mapping accuracy, (iii) analyse the synteny and collinearity between the genetic maps and the BRASUZ1 sequence in order to improve the first version of the *E. grandis* genome. Another key finding of this study was the discovery of segregation ratio distorting loci highlighting the presence of genes involved in pre- and/or post-zygotic reproductive barriers between these two species.

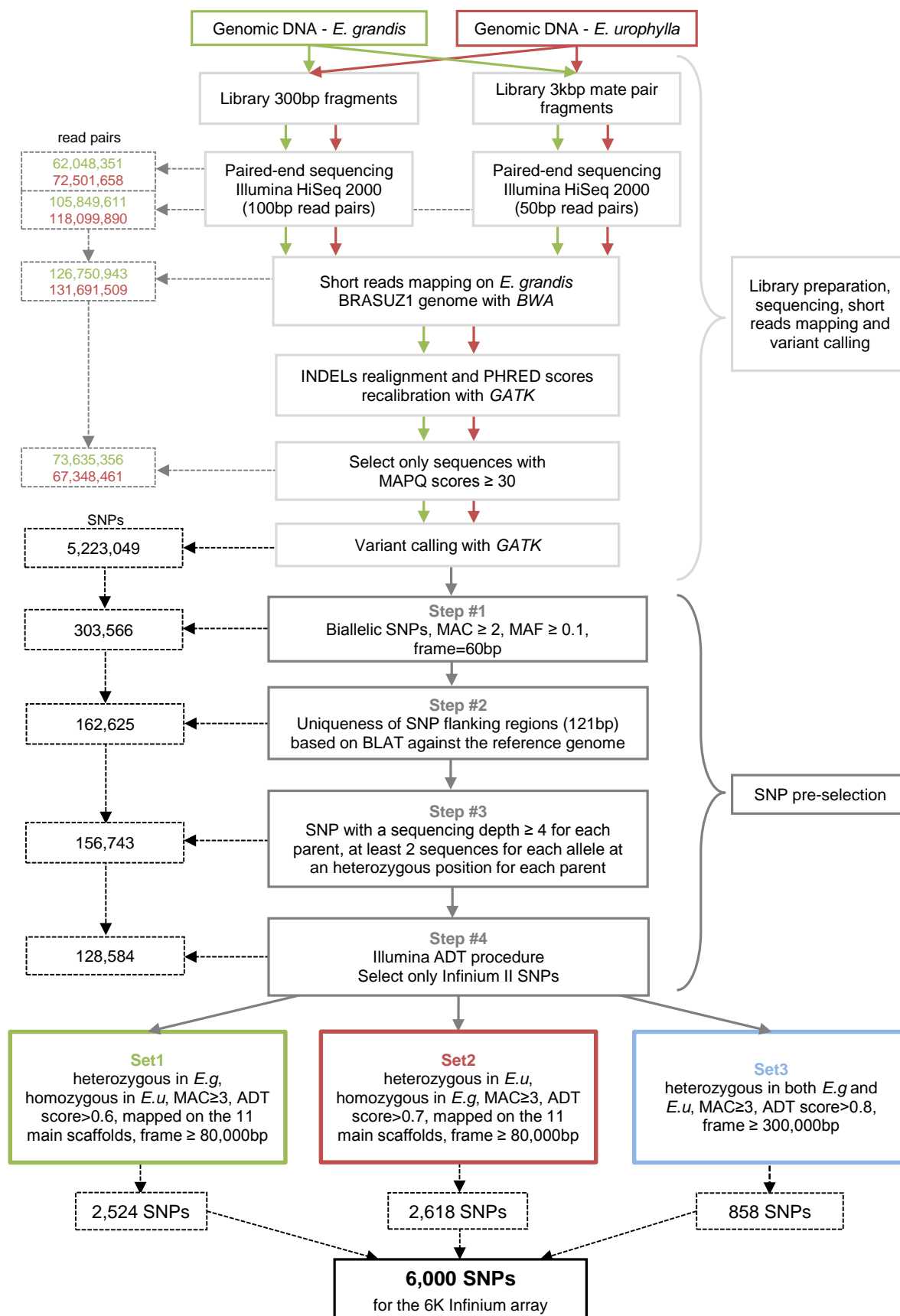


Figure 4.1 Flowchart for Infinium SNP array design.





## 4.2 Material and Methods

### 4.2.1 Plant material and DNA extraction

We used an interspecific cross between *E. urophylla* (*E.u.*, accession 14.144) as a female parent and *E. grandis* (*E.g.*, accession 9.21) as a male parent. These two species are phylogenetically close: they belong to the same subgenus (*Symphyomyrtus*) and the same section (*Latoangulatae*) (Steane *et al.*, 2002; 2011). The F1 mapping population totaled 2,700 full-sibs from which 1,025 were randomly selected for this study. For the genome sequencing of the two parents, total genomic DNA was extracted from frozen leaves using a modified protocol from Doyle and Doyle (1990). For SNP genotyping of the two parents and the 1,025 offspring, DNA was extracted from dry leaves with a protocol using magnetic beads (Smart D-N-Adem-Kit from Ademtech, Pessac, France). To assess DNA quality and quantity, the following analyses were carried out on each genotype: electrophoresis on 0.8% agarose gel and quantification by both spectrophotometry (NanoDrop 8000) and fluorescence (Quant-iT Picogreen).

### 4.2.2 SNP detection and array design

Whole genome resequencing was performed with an Illumina HiSeq2000 (GATC Biotech AG, Konstanz, Germany) for the two parental genotypes (*E.u.* and *E.g.*). Two libraries were constructed for each parent: (i) a 300 bp fragment size library (100 bp paired-end), and (ii) a 3kb mate pair fragment size library (50 bp paired-end). All steps from sequencing to array design are summarized in Figure 4.1.

The first version of the *Eucalyptus grandis* BRASUZ1 genome assembly (<http://www.phytozome.net/eucalyptus.php>) was used as the reference genome for mapping the short reads. Mapping was achieved independently for the two parental genotypes using the Burrows-Wheeler Alignment tool (BWA, version 0.6.1-r104, (Li & Durbin, 2009) with standard parameters. Then, the Genome Analysis Toolkit (GATK, version 1.5-21-g979a84a, (McKenna *et al.*, 2010) was used to realign insertion/deletion (INDEL) sites and recalibrate Phred quality scores (CountCovariates, IndelRealigner and TableRecalibration procedures with standard parameters). At the end of this step, one set of aligned short reads was available for each parent, combining 50 bp and 100 bp reads. Only sequences with high mapping quality (MAPQ score  $\geq 30$ ) were taken into account for the next steps.



The variant calling step was performed using GATK (UnifiedGenotyper procedure with options `–standcallconf 50.0` and `–standemitconf 10.0`) pulling together the two parental sets of aligned sequences. Different filters were applied to the initial set of SNPs through the variant calling procedure in order to obtain high quality *in silico* SNPs and design a 6K Infinium array (Figure 4.1). SNP pre-selection was divided into four steps. Firstly, only bi-allelic SNPs were conserved and INDEL variants were discarded. Selection of these SNPs was based on a minimum allele count ( $MAC \geq 2$ ), a minimum allele frequency ( $MAF \geq 0.1$ ) and a minimum distance of 60 bp between adjacent variations (SNP, INDEL) (Figure 4.1, Step#1). Secondly, the 121 bp sequence surrounding the SNP (i.e. the SNP and its two 60 bp flanking regions) was extracted for each SNP and checked for its uniqueness against the BRASUZ1 genome using BLAT software (Kent, 2002). SNP flanking sequences associated with more than one hit against the genome were discarded (Step#2). Then (Step#3), a more stringent selection was applied based on a home-made script: considering reads from both parents, only SNPs with a depth greater than 4 were kept. Moreover, each parental allele had to be present twice at the SNP position. Finally (Step#4), for all the conserved SNPs, a designability score was added using the Illumina Assay Design Tool (ADT) (Shen *et al.*, 2005). In order to maximize the number of SNPs on the array, only Infinium type II SNPs were considered.

Three categories of SNPs were finally selected based on their expected segregation patterns in the full-sib family (Set1, Set2 and Set3 as defined in Figure 4.1). Set1 consisted of SNPs heterozygous for the *E.g* parent and homozygote for the *E.u* parent (therefore segregating in the 1:1 Mendelian ratio), located on the main 11 scaffolds with the minimum distance between two consecutive SNPs of 80 kb, a  $MAC \geq 3$  and an ADT score  $> 0.6$ . Set2 comprised SNPs heterozygous only for the *E.u* parent and homozygote for the *E.g* parent (also segregating 1:1). Since a larger number of informative markers was initially available for *E.u*, the same criteria as for Set1 were used with a more stringent ADT score ( $> 0.7$ ). In Set3, SNPs were heterozygous in both parents (therefore segregating 1:2:1). No restriction on scaffold size was applied for Set3. With respect to the other criteria, we used a minimum distance between consecutive SNPs of 300 kb, a  $MAC \geq 3$  and an ADT score  $> 0.8$ . These three sets were used to design a 6K Infinium array (Illumina, San Diego, USA).



### 4.2.3 SNP genotyping

SNPs were genotyped by PEGASE-biosciences / Genes Diffusion (Douai, France). The results were analysed with GenomeStudio (Genotyping module V1.9, Illumina, San Diego, USA). Poorly performing individuals were removed from the analysis when they were lower than the following thresholds: 0.49 for the 10% GenCall score or 0.98 for the call rate. The clustering of each SNP was visually checked for its relevance to the expected inheritance pattern based on parental genotypes (Figure S4.1). When a cluster was not reliable, based on twenty four replicates of each parent, it was re-clustered manually. SNPs with a low fluorescence intensity, a call frequency lower than 0.90% or a GenTrain score lower than 0.4 were discarded. The reproducibility of the genotyping was evaluated using twenty-four replicates for each parent and two replicates for three offspring.

### 4.2.4 Linkage mapping

The genetic linkage analysis was performed with R software (Version 2.15.2; <http://cran.r-project.org/>) using the *onemap* V2.0-3 package (Margarido *et al.*, 2007; Mollinari *et al.*, 2009). Genetic maps were constructed for each parental genotype (*E.g* and *E.u*) according to a two-way pseudo-test-cross mapping strategy (Grattapaglia & Sederoff, 1994). The two parental maps were built independently with the same procedure. For linkage analysis we used SNPs and individuals with less than 4% and 2% of missing data, respectively. For the two parental maps, SNPs were grouped into linkage groups (LGs) with a stringent threshold (LOD score  $\geq 30$ ). In a first step, marker ordering within linkage groups was performed with the RECORD algorithm (Os *et al.*, 2005) implemented in *onemap* using only SNPs segregating 1:1 (Set1 and Set2, test-cross markers) to build the framework maps. In the second step, SNPs segregating 1:2:1 (Set3, inter-cross markers) were added to the framework maps using the RECORD algorithm to build complete maps. Genetic distances (cM) were calculated using the Haldane mapping function (Haldane, 1919). SNP segregations were tested for goodness-of-fit to the expected Mendelian segregation ratios using  $\chi^2$  tests with the level of significance adjusted for simultaneous multiple tests (Benjamini & Yekutieli, 2001) within each LG of the parental maps.



#### 4.2.5 Linkage mapping accuracy

The effect of the sample size (number of offspring) and marker density on linkage map accuracy was tested based on re-sampling in the whole dataset. Three sample sizes (100, 200 and 500 individuals) and four marker densities (using only markers segregating 1:1) were used. For each sample size, a random draw (1,000 times) was performed from the 1,020 offspring. Marker selection was based on genetic distances obtained with the framework maps. Four density classes were selected: one marker every 5, 2.5 and 1.2 cM and all markers (referred to as MD1, MD2, MD3 and MD4 respectively). In total, 12,000 genetic maps for each parent (3 sample sizes  $\times$  1,000 samples  $\times$  4 marker densities) were built using the RECORD algorithm of *onemap*. For each map, grouping into linkage groups was performed using the framework maps as a priori information.

### 4.3 Results

#### 4.3.1 Resequencing, SNP selection and SNP-array design

Considering all libraries, resequencing of the two parental genotypes resulted in 49.3 Gb of sequences. A total of 62,048,351 (72,501,658) 100 bp read pairs and 105,849,611 (118,099,890) 50 bp read pairs were obtained for *E.g* (*E.u*). According to the estimated physical size of the *E.g* (640 Mb) and *E.u* (650 Mb) genomes (Grattapaglia & Bradshaw Jr., 1994), the theoretical mean haploid genome coverage was 35.9X and 40.5X for *E.g* and *E.u*, respectively. The proportion of aligned reads on the BRASUZ1 genome was higher for *E.g* (75.5%) than for *E.u* (69.1%), as might be expected given the species of the reference genome (*E. grandis*) and the small phylogenetic distance between *E.g* and *E.u*. The distribution of aligned reads along the genome was homogenous between scaffolds and between read types: 100 bp or 50 bp (Figure S4.2). Only robustly aligned reads (mapQ >30) were used for SNP detection, i.e. 51.1% (*E.g*) and 58.1% (*E.u*) of the aligned reads.

The variant calling procedure with GATK resulted in 5,223,049 SNPs being either polymorphic within or between the two parental genotypes. After applying stringent selection criteria to the initial set of detected SNPs (Figure 4.1, Step#1 to #4), 128,584 high-quality *in silico* SNPs were obtained. Lastly, we selected three complementary sets of SNPs based on their comprehensive distribution along the genome and on their expected segregation patterns in the progeny (Figure S4.3), as follows: Set1 presented 2,524 SNPs heterozygous in *E.g* only



(1:1 segregation), Set2 presented 2,618 SNPs heterozygous in *E.u* only (1:1 segregation) and Set3 presented 858 SNPs heterozygous in both parents (1:2:1 segregation). The selected SNPs were evenly spaced along the BRASUZ1 genome with an average physical distance between adjacent SNPs of 100 kb ( $\pm$  163 kb). However, some regions amounting to 8% of the scaffold length on average were less covered due to a distance of more than 1 Mb between adjacent markers. Most of the time these regions were characterized by a high rate of missing data (gaps) on the reference genome (e.g. 25% from 20 Mb to 22 Mb and 23% from 32 Mb to 34 Mb on scaffolds 3 and 5, respectively, see Figure S4.3).

Table 4.1 Genotyping results: distribution of polymorphic, monomorphic and failed (i.e. no call) SNPs for the three sets presented in the material and methods section: set 1 and 2 testcross markers informative for the *E. grandis* and *E. urophylla* parents, respectively; set 3 inter-cross markers informative in both parents.

	Polymorphic	Monomorphic	No call	Total
Set 1	1,849 (88.6%)	82 (3.9%)	157 (7.5%)	2,088
Set 2	1,793 (78%)	366 (15.9%)	139 (6%)	2,298
Set 3	725 (91.9%)	21 (2.7%)	43 (5.4%)	789
Total	4,367 (84.4%)	469 (9.1%)	339 (6.6%)	5,175

#### 4.3.2 Genotyping quality

The success rate for custom BeadChip manufacturing was 86.25%, i.e. 5,175 SNPs were available for genotyping. After SNP clustering and quality control with GenomeStudio, five genotypes with a low call rate ( $< 0.98$ ) or a low 10% GenCall ( $< 0.49$ ) were removed from further analysis. In all, 339 SNPs with technical failures (i.e. no call) were discarded (Table 4.1). Of the remaining 4,836 SNPs, 469 (9.7%) were monomorphic and 4,367 (90.3%) exhibited clear segregation within the full-sib family. The proportion of monomorphic markers was lower for Set1 (3.9%) compared to Set2 (15.9%). This must have been related to the differences in genome structure between the two species and the less optimum alignment of *E.u* short reads on the *E. grandis* BRASUZ1 genome. Moreover, the distribution of monomorphic markers between scaffolds was different between the two parental sets (Table S4.3). Within scaffolds, some regions from 1 Mb to 3.9 Mb were found to gather

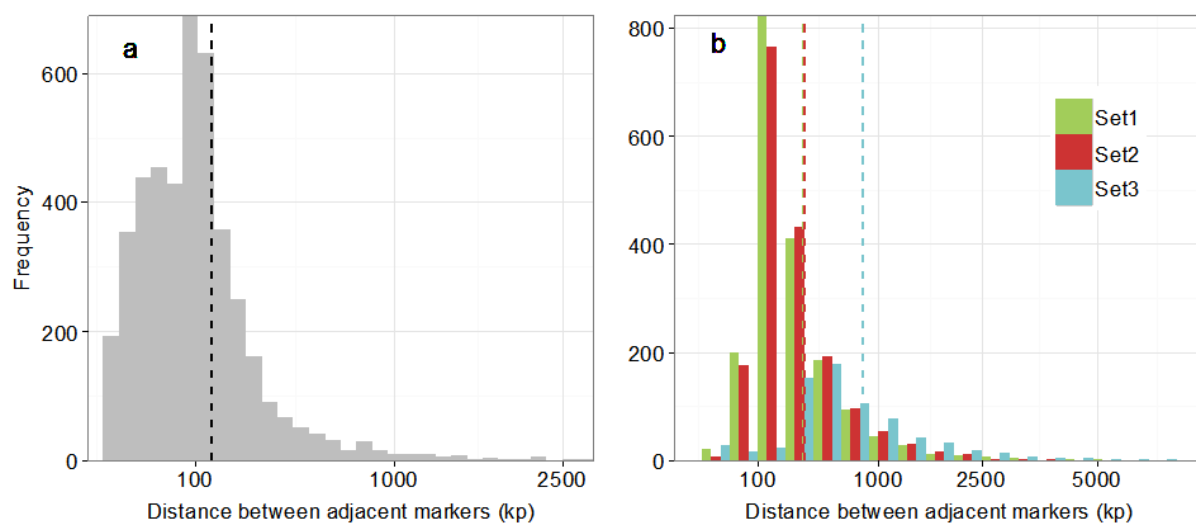


Figure 4.2 Distribution of inter-marker distances for the polymorphic SNPs for a/ the whole dataset and b/ the three SNP sets. The vertical dashed lines represent the mean distance between adjacent markers.

Table 4.2 Characteristics of framework (test-cross markers) and complete linkage maps (test-cross and inter-cross markers) for *Eucalyptus grandis* (E.g) and *E. urophylla* (E.u).

	Framework map		Complete map	
	<i>E.g</i>	<i>E.u</i>	<i>E.g</i>	<i>E.u</i>
Map length (cM)	821.66	885.92	912.59	903.99
Number of SNPs ( total)	1,832	1,773	2,551	2,491
Number of SNPs (1:1)	1,832	1,773	1,832	1,773
Number of SNPs (1:2:1)	-	-	719	718
Number of bins	1,429	1,353	1,904	1,799
Distance between SNPs (cM)	0.454	0.498	0.358	0.362
Distance between bins (cM)	0.578	0.652	0.48	0.5
Distorted SNPs (%)	21	30.7	19.8	28.7
Linkage map coverage (%)	98.2	97.2	98.3	97.3

Bin: unique loci on a genetic map



monomorphic SNPs for one parent and polymorphic for the other: on scaffold 6 for *E.g* and on scaffolds 4, 5, 6 and 10 for *E.u* (Figure S4.4). For the polymorphic SNPs, the average call frequency was 99.8% ( $\pm 0.58\%$ ), the average GenTrain score was 79.9% ( $\pm 6.3\%$ ) and the average cluster separation score was 0.91 ( $\pm 0.18$ ). Considering the three SNP sets, the average physical distance between adjacent polymorphic SNPs was 138 kb (Figure 4.2A). Taking the set separately, the average inter-marker distance was larger: 321 kb for Set1, 332 kb for Set2 and 797 kb for Set3 (Figure 4.2B). Genotyping reproducibility over all polymorphic SNPs was at least 99.95%. Using version 1.1 of the predicted gene models of the BRASUZ1 genome, 38% of the polymorphic SNPs were found to map into genes (Set1: 32%, Set2: 39.5% and Set3: 52.1%). Moreover, 78% of the polymorphic SNPs were located within 5 kbp of predicted gene models.

### 4.3.3 Genetic Mapping

#### 4.3.3.1 Construction of high-resolution genetic linkage maps

The maps were established using the same procedure and independently for the two parents. Set1 and Set2 (SNPs segregating 1:1) were first used for the construction of two framework maps. The total map length was lower for *E.g* (821.7 cM) than for *E.u* (885.9 cM) (Table 4.2, Table S4.5). Framework maps provided a high resolution with one marker every 0.45 cM (*E.g*) and 0.50 cM (*E.u*) on average. Despite greater coverage of the reference genome for *E.g* maps (98.2%) than for *E.u* maps (97.2%), the *E.g* map was significantly shorter (821 cM) compared to that of *E.u* (885 cM). Moreover, this difference was underestimated because of the low coverage of scaffold 4 for *E.u* (85.6%) resulting in a shorter LG4 length (5.5 cM) compared to *E.g*. Overall, no particular trend in terms of LG length was found between species (Table S4.5). SNPs with significant segregation distortion ( $P_{\text{adjust}} < 0.01$ ) were kept for linkage map construction and accounted for 21% and 30.7% of the mapped SNPs for *E.g* and *E.u*, respectively (Table 4.2). SNPs with segregation distortion occurred in large regions (SDRs) on several LGs with differences between parental maps (Figure S4.5). LG1 and 3 presented a high level of distortion in *E.g* only, LG5, 6, 7 and 11 only in *E.u*, LG2 presented a high level of distortion in both parental maps and LG4, 8, 9 and 10 presented no distortion in either parental map (Figure S4.5).

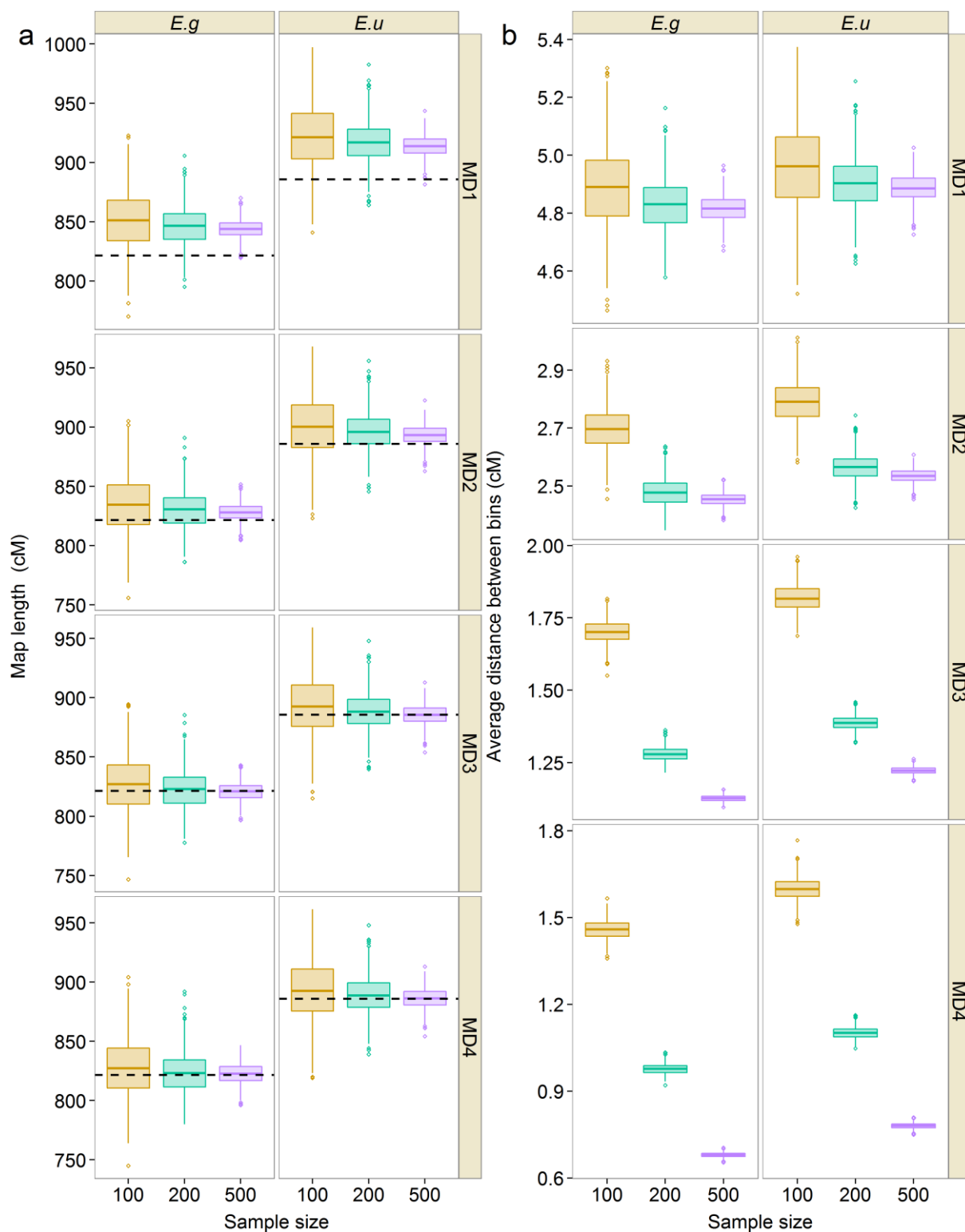


Figure 4.3 Boxplot of map length (a) and average distance between bins (b) for different combinations of marker density (MD) and sample size resulting from 1,000 maps for each combination. The horizontal dashed lines represent the length of the framework maps for *E. grandis* (*E.g*) and *E. urophylla* (*E.u*).





In a second step, 719 SNPs from Set3 (segregating 1:2:1) were added to the framework maps. The resulting maps comprised 2,552 and 2,493 SNPs for a total map length of 912.6 cM for *E.g* and 904 cM for *E.u* (Table 4.2, Table S4.6), corresponding to an additional 90.9 and 18.1 cM compared to the framework maps, for *E.g* and *E.u* respectively. The resolution provided by these two maps was high with one marker every 0.36 cM. Conversely to the comparison between parental framework maps, the total map length was slightly greater for *E.g* (8.3 cM longer) than for *E.u*, probably due to less optimum coverage of scaffold 4 (85.6%) for *E.u*. Thus, the integration of SNPs segregating in both parents (1:2:1) tended to hide parental differences in map length observed for the framework maps based 1:1 segregating SNPs alone. However, the high density of 1:2:1 SNPs every 1.26 cM (*E.g*) and 1.34 cM (*E.u*) enabled a clear identification of orthologous regions between *E.g* and *E.u* parents. The order of common SNPs between parental maps was well conserved, except for 9.3% of inverted SNPs mainly located on LG 2, 3, 5 and 7 (Figure S4.6). These inversions of 0.4 cM on average may have been caused by a decrease in precision in the recombination rate estimations between pairs of markers segregating 1:1 and 1:2:1 (Ritter *et al.*, 1990). Indeed, between test-cross and inter-cross markers, only half of the genotypes are informative for linkage analysis.

#### 4.3.3.2 Effect of sample size on mapping accuracy

The effect of sample size and marker density (MD) on genetic mapping accuracy was tested using a random draw in the whole dataset. Classes of MD were based on framework maps (Table S4.7). The results obtained were very similar for both parental maps. For a given marker density, the sample size had little impact on the average map length (Figure 4.3A, Table S4.8), with a general trend towards longer map lengths for small sample sizes (e.g. for MD1 the average map length (*E.g*) was 851 cM, 846.2 cM and 844 cM for  $n = 100$ , 200 and 500, respectively). As marker density increased from MD1 to MD4 the average map length for all sample sizes tended to reach that of the framework maps with a difference of 22.3 cM (MD1,  $n=500$ ) and 1cM (MD4,  $n=500$ ) for *E.g*. The combined effect of sample size and marker density was more significant on the average distance between bins (unique loci on a

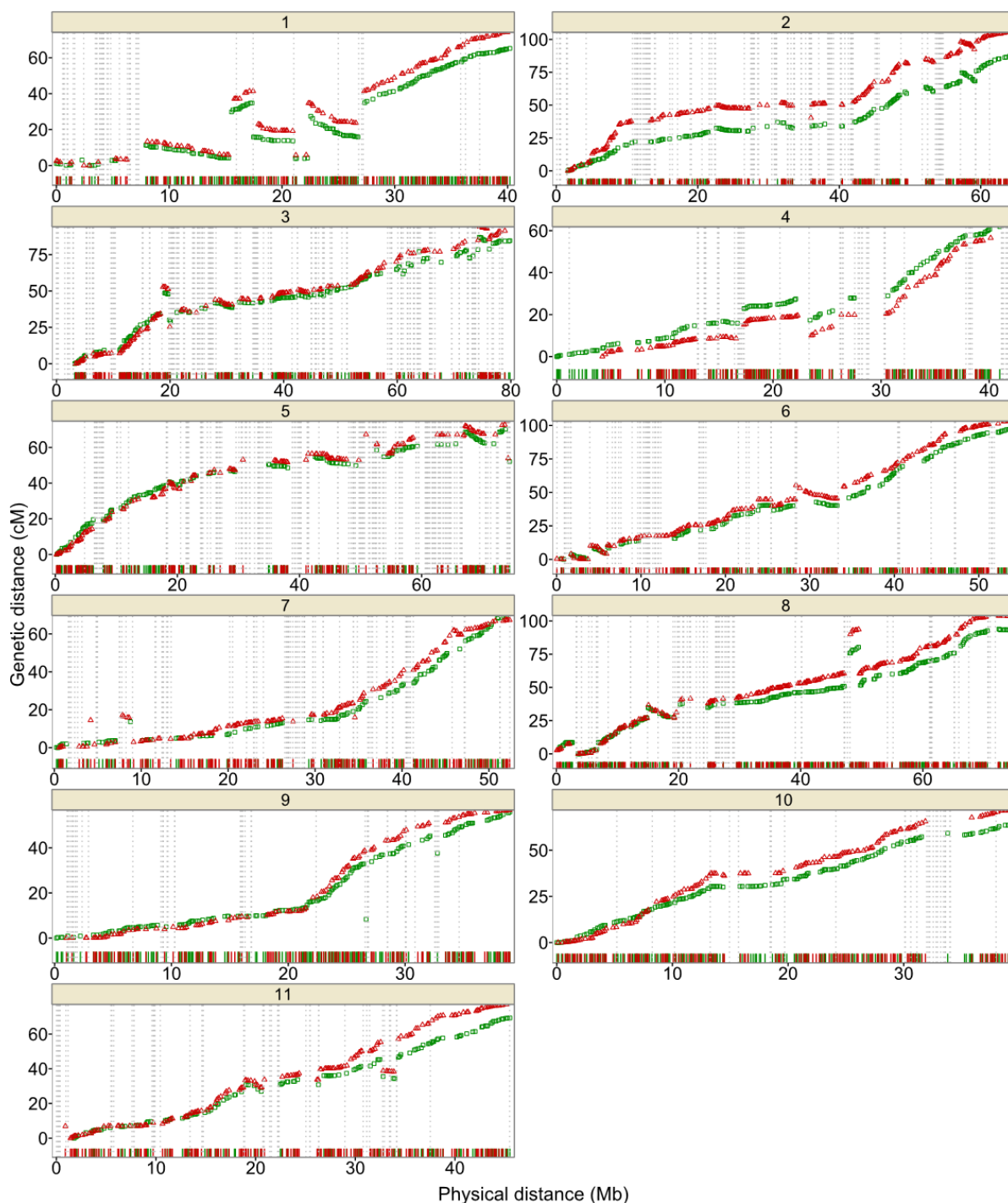


Figure 4.4 Physical position (in Mb) and the genetic location (in cM) for the SNPs mapped on *E. grandis* (green) and *E. urophylla* (red) framework maps. Vertical dotted lines in grey indicate gaps larger than 10 kb on the BRASUZ1 genome sequence. Only markers with a conserved chromosomal assignment between genetic and physical maps are represented.



genetic map (Figure 4.3B, Table S4.8). Indeed, for low marker density (MD1) the sample size had no effect on the average distances between bins: 4.9, 4.8 and 4.8 cM for  $n=100$ , 200 and 500, respectively. On the other hand, for a higher marker density (MD4), the increase in sample size improved the resolution of the maps (e.g. for *E.g* the average distances between bins were 1.5, 1 and 0.7 cM for  $n=100$ , 200 and 500, respectively).

#### 4.3.4 Comparison between genetic maps and the BRASUZ1 genome

The high resolution in marker ordering made it possible to compare SNP locations on the *E.g* and *E.u* framework maps with their respective position on the reference genome on a fine scale (Figure 4.4). Overall, the two high-resolution linkage maps displayed a high level of collinearity with the reference genome, i.e. conserved order of SNPs between their genetic and physical locations. These collinear regions amounted to 85.7% of the 605.9 Mb of the reference genome (i.e. the main 11 scaffolds) for *E.g* and 84.1% for *E.u*.

The high-resolution framework maps also highlighted discordant genetic regions (a genetic region is defined by at least two successive mapped SNPs) in comparison to the genome assembly. We distinguished (i) differences in region ordering and/or orientation within chromosomes called non-collinear regions (NCR), and (ii) non-conserved chromosomal assignments between genetic and physical maps, called non-syntenic regions (NSR). In all, 49 and 51 NCRs were identified for *E.g* and *E.u*, representing 13.9% and 15.3% of the reference genome, respectively (Table 4.3 and Figure 4.5). Of the 57 NCRs (Table S4.9), 43 concerned 9 out of the 11 scaffolds, and corroborated between the two parental maps. They represented 13.5% (*E.g*) and 14.4% (*E.u*) of the genome size. Considering only these common NCRs, scaffold 1 presented the largest differences between physical and genetic maps accounting for 46.4% (*E.g*) and 43.3% (*E.u*) of the scaffold size (Table S4.9). The second scaffold with the largest proportion of common NCRs was scaffold 5 with 22.8% (*E.g*) and 27.1% (*E.u*) of its size. Specific NCRs were also found for *E.g* (6) and *E.u* (8) and were mainly related to a small number (or the absence) of SNPs in one parent compared to the other (Table S5). These specific NCRs were distributed over scaffolds 1, 3 and 7 for *E.g* and scaffolds 1, 2, 3, 6, 7 and 9 for *E.u*. For both parental maps scaffold 7 only presented specific NCRs.

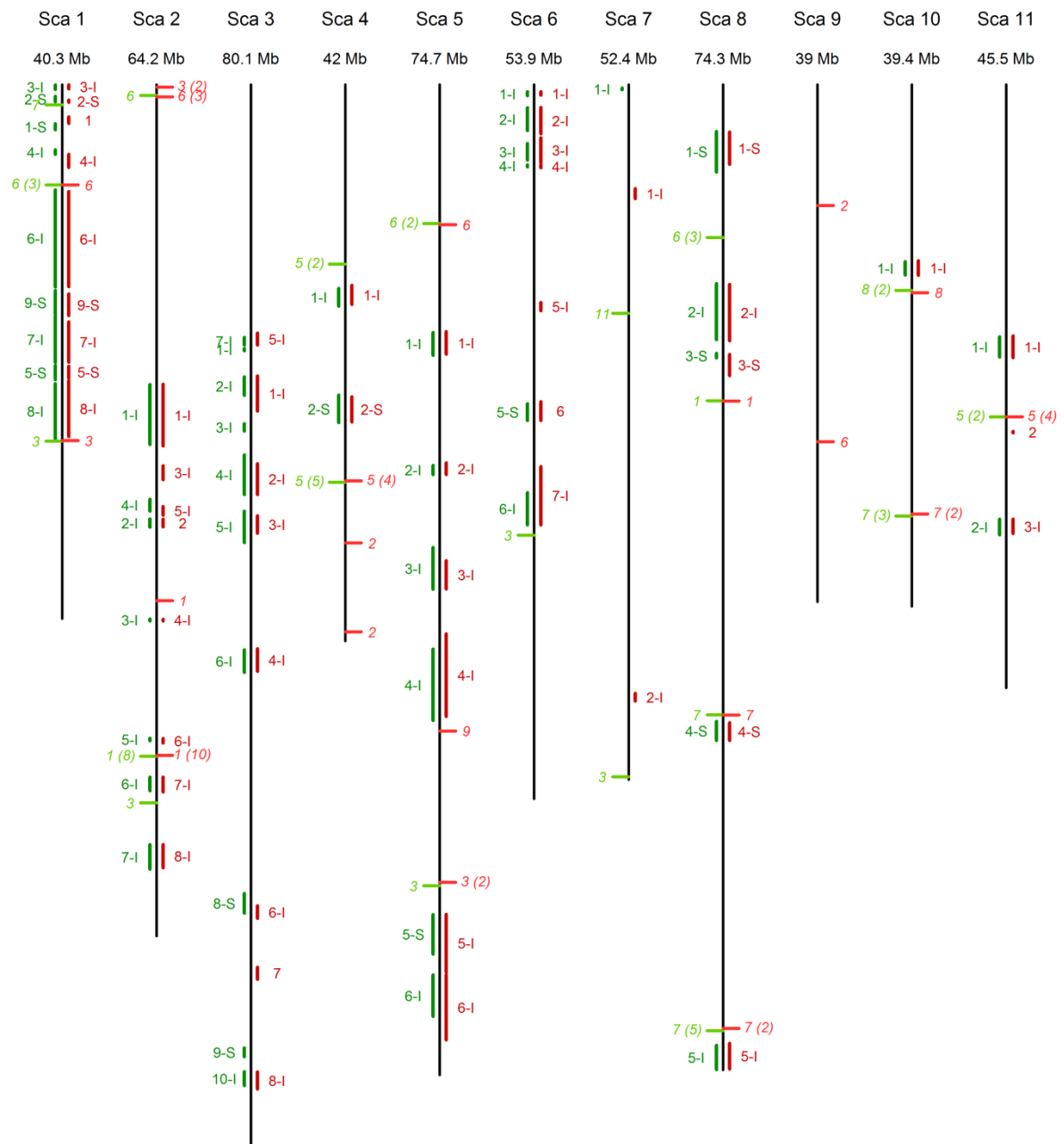


Figure 4.5 Physical location of non-collinear regions (NCRs) and non-syntenic regions (NSRs) on the *BRASUZI* genome sequence. Vertical black lines represent the physical skeleton of the main eleven scaffolds (Sca). Discordant regions between physical and genetic maps are in green for *E. grandis* (left) and red for *E. urophylla* (right). Vertical lines on each scaffold represent NCRs. They are numbered in the same way as in Table S8. “I” indicates inverted NCRs while “S” indicates displaced NCRs. Horizontal tick bars indicate NSRs, followed by the linkage group number onto which the NSR is mapped. The number in brackets indicates the number of markers involved in the NSR (if greater than 1).



Table 4.3 Features of the non-collinear region (NCR) and non-syntenic region (NSR) between genetic framework maps, *E.g*: *E. grandis* and *E.u*: *E. urophylla*, and the BRASUZ1 genome sequence. Common NCRs or NSRs are supported by both parental maps.

		E.g	E.u
NCR	All	Number of regions	49
		Number of SNPs	393
		% of SNPs	21.5
		% of genome size	13.9
	Common	Number of regions	43
		% of regions	87.8
		Number of SNPs	378
		% of SNPs	20.6
		% of genome size	13.5
			14.4
NSR	All	Number of regions	20
		Number of SNPs	45
		% of SNPs	2.5
		% of genome size	1
	Common	LG final	1, 3, 5, 6, 7, 8, 11
			1, 2, 3, 5, 6, 7, 8, 9
		Number of regions	13
		% of regions	65
		Number of SNPs	35
		% of SNPs	1.9
		% of genome size	0.8
		LG final	1, 3, 5, 6, 7, 8

To a lesser extent, non-conserved chromosomal assignments (NSR) were revealed by the comparison between the physical and genetic maps (Figure 4.5 and Table 4.3). Except for scaffold 3, we identified 27 NSRs on all the other scaffolds that represented 1% (*E.g*) and 0.7% (*E.u*) of the targeted scaffold size, of which 13 were common to the two parental maps (Table S4.10). Considering common NSRs, scaffold 2 gathered the largest NSR with 2.9% (*E.g*) and 3.5% (*E.u*) of the scaffold size. Specific NSRs were also found on scaffolds 3, 6, 7 and 9. As explained previously for NCRs, specific NSRs were mainly related to the absence of SNPs in the other parental map, but in four cases (on scaffolds 6, 7 and 8 for *E.g* and on scaffold 9 for *E.u*) the closest marker in the orthologous map was well localized (Table S4.10). This could be explained by the presence of paralogous SNPs. The NCRs and NSRs that were corroborated by the two parental genetic maps strongly suggest miss-assembly in the BRASUZ1 genome rather than structural rearrangements in the two genotypes studied here.

Besides the comparison of SNP order, 13 additional scaffolds totalizing 6.9 Mb (8.1% of the 85 Mb of the 4,941 unmapped scaffolds) were also mapped on LGs, using one SNP per



scaffold (Table S4.11). Of these 13 additional scaffolds located on 7 LGs, 3 were positioned on LG2 and 3, 2 on LG4 and 8, and 1 on LG1, 5 and 9. All additional scaffolds were located at a similar genetic position on both parental maps (Table S4.11) validating their position.

## 4.4 Discussion

### 4.4.1 Performance of the SNP Array

This study is the first to report on the designing of an Infinium SNP array for *Eucalyptus*. Our approach for SNP detection was based on whole genome resequencing of the two parental genotypes of a mapping population with high sequencing depth ( $\approx 40\times$ ). Stringent selection criteria resulted in a high genotyping success rate (SR), i.e. 84.4% of the genotyped SNPs were polymorphic in the studied pedigree. Such a high SR was also reported in *Populus trichocarpa* using a similar resequencing approach (85.6% in (Slavov *et al.*, 2012) and 93.1% in (Geraldes *et al.*, 2013)). Chagné *et al.* (2012) on *Malus pumila*, reported a lower SR of 72.2% despite an additional validation step on a subset of SNPs with the GoldenGate technology. This contrast with the lower values is generally obtained in non-model forest tree species using the Infinium technology and largely based on SNPs detected from RNA-seq data. For instance, Pavy *et al.* (2013) designed two SNP arrays in *Picea glauca*, and reported SRs of 55.8% and 67.6%. Along the same lines, Howe *et al.* (2013) obtained a SR of 72.5% in *Pseudotsuga menziesii*. This disparity between studies might be explained by different factors such as the nature of the primary data (RNA vs whole genome DNA) and the stringency of thresholds at different stages of SNP detection and selection.

### 4.4.2 Linkage mapping accuracy

The two framework genetic linkage maps were built independently with SNPs segregating 1:1. Contrary to other high density mapping studies in *Eucalyptus* (Hudson *et al.*, 2012a; Kullán *et al.*, 2012; Petroli *et al.*, 2012), we did not combine parental maps in a consensus map because of the loss of parental specific feature and the loss of accuracy introduced by combining markers with different segregation ratios, resulting in a loss of information (Ritter *et al.* 1990). The use of markers with different segregation types (mandatory for combining parental maps) indeed reduces mapping accuracy, as illustrated by the inversion of 9.3% of inter-cross markers between the two parental maps obtained using both test-cross and inter-cross markers. Additional factors could also lead to poorer order and thereby bias the



resolution of a genetic map, e.g. inconsistencies between individual maps such as local reordering and/or large displacements (Jackson *et al.*, 2008; Wu *et al.*, 2011; Ronin *et al.*, 2012).

The two parental framework maps provided a comprehensive view of the *Eucalyptus* genome with 98.2% and 97.2% coverage of the main 11 scaffolds and an average inter-marker distance of 321 kb and 332 kb for *E.g* and *E.u*, respectively. The corresponding map lengths were 821.7 cM and 885.9 cM with a resolution of one SNP every 0.45 cM and 0.50 cM for *E.g* and *E.u*, respectively. Previous genetic size estimates of the genome in different *Eucalyptus* species ranged from 632 to 1,815 cM (Table S4.1) with the lower and higher limits obtained with low density and/or non-saturated genetic maps (Brondani *et al.*, 2006; Rocha *et al.*, 2007). Compared with recent studies (Hudson *et al.*, 2012a; 2012b; Kullán *et al.*, 2012; Petroli *et al.*, 2012) which used a DArT array to build dense genetic maps, our estimates of the *Eucalyptus* genome size showed a reduction of at least 20% (*E.g*) and 15% (*E.u*). The size of a genetic map characterizes the number of recombination events per meiosis captured in the mapping population (Crismani *et al.*, 2013). In our study the number of recombination events was higher than the most recent ultra-dense genetic maps in *Eucalyptus* (Table S1). Moreover, these dense genetic maps were consensus maps which could artificially inflate total map length. Thus, the two parental framework maps based on robust data only provided the most accurate estimates so far of the genetic map length for *E. grandis* and *E. urophylla*.

To confirm the robustness of the two parental genetic maps generated in this study, we compared the results obtained with RECORD to those obtained with Joinmap, one of the most widely used softwares for the estimation of genetic maps (Van Ooijen, 2011). The total map lengths were similar with 821.1 and 821.7 cM for *E.g* and 884.9 and 885.9 cM for *E.u* for JoinMap and RECORD, respectively. The SNP order was very similar with 99.2% and 99.6% of SNPs exhibiting the same order between the two algorithms for *E.g* and *E.u*, respectively (Table S4.5). Inversions only occurred with tightly linked SNPs with a maximum distance between inverted SNPs of 0.1 cM, confirming the robustness of the framework maps. Moreover, these results were corroborated by highly reliable orders obtained with the resampling of F1 offspring with up to 0.7% of inverted markers even with a small sample size (100).





#### 4.4.3 Validation and improvement of the *Eucalyptus* BRASUZ1 genome structure

The use of two independent high-resolution genetic maps enabled cross-validation in the comparison between genetic maps and the reference genome sequence. Our results confirmed the robust assembly of the BRASUZ1 genome. Indeed, collinear regions accounted for approximately 85% of the reference genome. This first version was assembled using a DArT genetic map (Kullan *et al.*, 2012) and independently validated by Petroli *et al.* (2012). If inconsistencies between physical and genetic order were reported for different LGs (e.g. 1 and 4), they were attributed to poor marker quality rather than wrong scaffold assembly (Petroli *et al.*, 2012). The common NCRs and NSRs between parental maps which accounted for nearly 14% and 1%, respectively, suggest the contrary and advocate an improved version of the BRASUZ1 genome assembly. In addition to the results relative to the organization of the main 11 scaffolds, 13 additional scaffolds totalizing 6.9 Mb were positioned and validated by both parental maps. As suggested by Petroli *et al.* (2012) the majority of small unanchored scaffolds (<20 kb) must correspond to an already assembled part of the genome (i.e. alternative haplotypes). Indeed the BRASUZ1 genotype was produced after one generation of selfing leaving half of its genome in the heterozygous state in theory. If this hypothesis is true, these additional 13 scaffolds of medium to large size (0.06 to 1.4 Mb) could be a real missing part of the genome. Thus, their integration could improve the completeness of the BRASUZ1 genome.

The comparison between parental genetics maps and the reference genome highlighted the high conservation of genome structure between the two species (*E.g* and *E.u*). Although differences were highlighted by genotyping (e.g. LG4, 3.9Mb with monomorphic markers for *E.u*), no inconsistencies involving a large number of SNPs were found between the two maps. This collinearity and synteny between *Eucalyptus* species had already been reported (Myburg *et al.*, 2003; Hudson *et al.*, 2012b) and is consistent with the small phylogenetic distance between these two interfertile species of the *Symphyomyrtus* subgenus (Potts & Dungey, 2004). Moreover, their genomes were found to be close in terms of physical size with a larger genome for *E.u* (Grattapaglia & Bradshaw Jr., 1994; Praça *et al.*, 2009).





#### 4.4.4 Distribution and effect of SDRs

Linkage mapping analysis was carried out with distorted SNPs, not only to maintain high genome coverage in such regions of the genome but also because segregation distortion has little or no effect on mapping accuracy (Hackett & Broadfoot, 2003; Hudson *et al.*, 2012b), which was confirmed in our study by the same marker order found between *E.g* and *E.u* LGs displaying different levels of segregation distortion between parental maps (e.g. LG1 and 11). Moreover, the same marker order was found between different samples of individuals with or without segregation distortion in a given LG (data not show). The number of offspring (1,025) and the high synteny between physical map and parental genetic maps made us confident in the nature of the distortion, i.e. evidence for biological causes compared to bias in the genotyping process. Our study therefore offers a precise landscape of segregation distortion along the eucalyptus genome.

Distorted SNPs occurred in large segments (so called SDRs) and accounted for 21% (*E.g*) and 30.9% (*E.u*) of the framework maps. Large differences in segregation distortion levels between parental maps were found on six LGs (1, 3, 5, 6, 7 and 11). These differences between parental maps could be due to pre-zygotic events. Similar distortion patterns (e.g. in LG2) could be related to post-zygotic events. SDRs could be related to different physiological and genetic factors occurring during the reproductive process, such as pollen-tube competition (Arnold *et al.*, 1993; Rahmé *et al.*, 2009; Zhang *et al.*, 2011), pollen-pistil incompatibility, which was found to increase with phylogenetic distance in *Eucalyptus* species (Gore *et al.*, 1990; Ellis *et al.*, 1991) and in the closely related *Corymbia* species (Dickinson *et al.*, 2012), negative epistatic interactions among alleles (Törjék *et al.*, 2006; Bikard *et al.*, 2009) or lethal genes in a homozygous state. Hybridisation enhances some of these processes (Rieseberg & Carney, 1998) thus accentuating the level of segregation distortion in hybrid pedigrees, as found in genetic mapping studies in *Eucalyptus* (Table S4.1) or in *Populus* (Bradshaw & Stettler, 1994; Yin *et al.*, 2004; Yin *et al.*, 2008). In *Eucalyptus*, segregation distortion has been reported in similar proportions in interspecific crosses (Myburg *et al.*, 2003; Brondani *et al.*, 2006; Kullán *et al.*, 2012) and to a lesser extent in intraspecific crosses (Thamarus *et al.*, 2002; Hudson *et al.*, 2012b) suggesting that genes involved in reproductive barriers underlined these regions. As the success of tree breeding could be significantly affected by these genes, we list the genes at all the segregation distortion peaks (i.e. maximum significant value, Table S4.12).



## 4.5 Conclusion and prospects

Our SNP array maximizing genome coverage with 6,000 SNPs evenly spaced along the main 11 scaffolds of the BRASUZ1 genome resulted in: (i) 84.4% of polymorphic SNPs segregating in 1,025 full-sibs, (ii) two high-resolution genetic linkage maps, (iii) an improved version of the *Eucalyptus* genome assembly, and iv/ the discovery of genes possibly involved in segregation distortion in this cross. With this SNP-array, it should be possible to capture LD at long physical distances in small sized structured populations and thereby enable genomic selection to be carried out in the eucalyptus breeding programme that CIRAD and CRDPI are implementing in the Republic of Congo. In addition, the size of the full-sib family planted out in the field in 2010 with 2,700 offspring will enable fine-mapping of a QTL-rich region for lignin content and composition (Gion *et al.*, 2011).

### Authors' contributions

JB: DNA extraction. AM, EM and JB: Field sampling. EM, CK, NA: Array design. JB: Genotyping analysis and linkage analysis. JB, JMG, CP and EM: manuscript preparation. JMG: Funding and overall supervision. All authors read and approved the final manuscript.

### Acknowledgments

The authors would like to thank CRDPI (*Centre de Recherches sur la Durabilité des Plantations Industrielles located in the Republic of the Congo*). The authors also thank Matthieu Falque who provided the R script for Fig 5. This study was supported by grants from FEDER (ABIOGEN project, no. Presage 32973) and ERANET Plant KBBE (ANR(FR)-10-KBBE-0007 “34Joule”) JB received a PhD fellowship from the BIOS department of CIRAD (*Centre de coopération Internationale en Recherche Agronomique pour le Développement*).



## Supplemental figures and tables

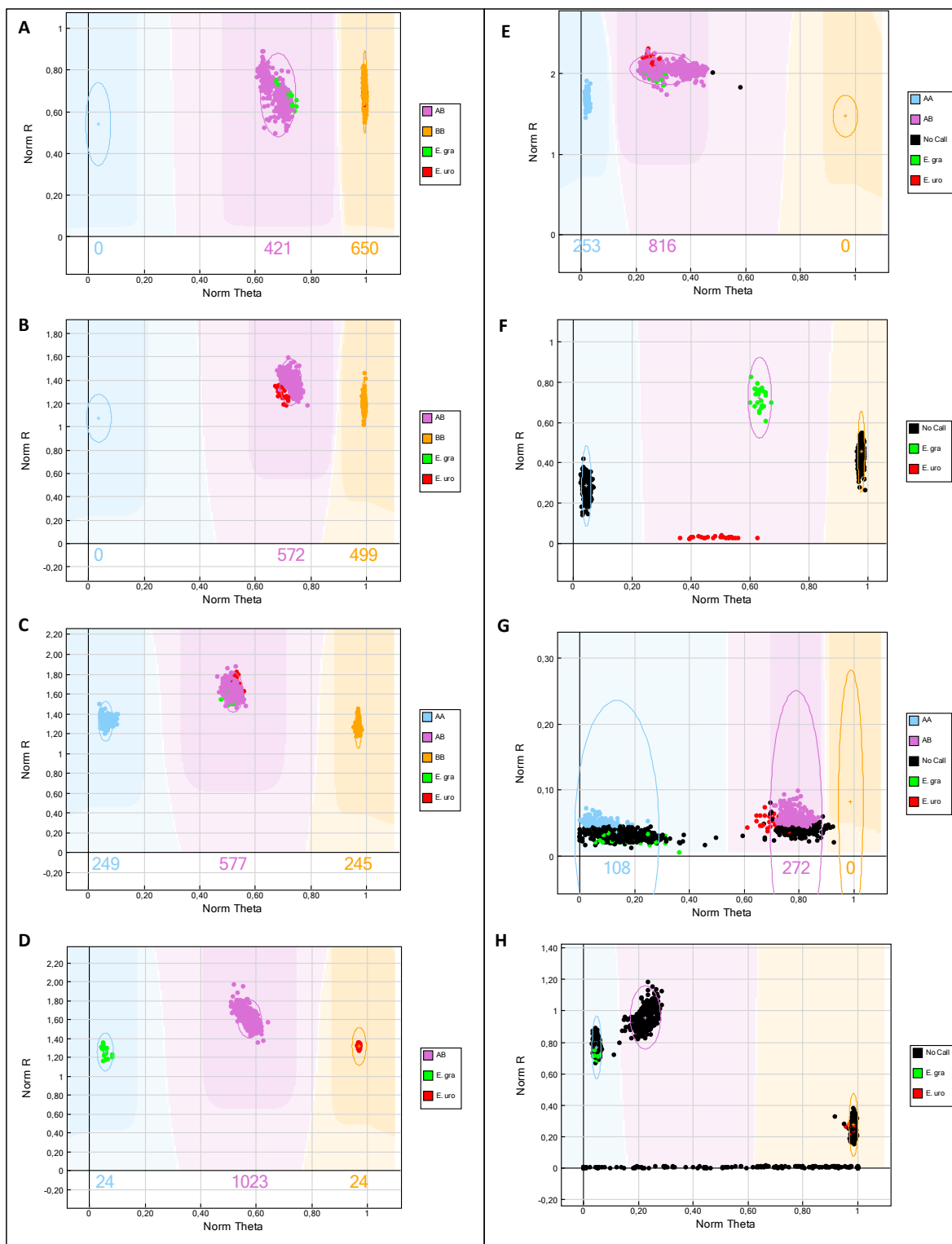


Figure S4.1 Examples of SNP clusters obtained with GenomeStudio software. Accurate genotype calling for polymorphic (set1 in the material and methods section (panel A), set2 (B) and set3 (C)) and monomorphic SNPs (D). SNPs with an unexpected segregation pattern (E, F) or low intensity (G, H) considered as failed assays. Twenty-four replications for each parental genotype were available and are represented as green dots for *E. grandis* (E.g) and red dots for *E. urophylla* (E.u).

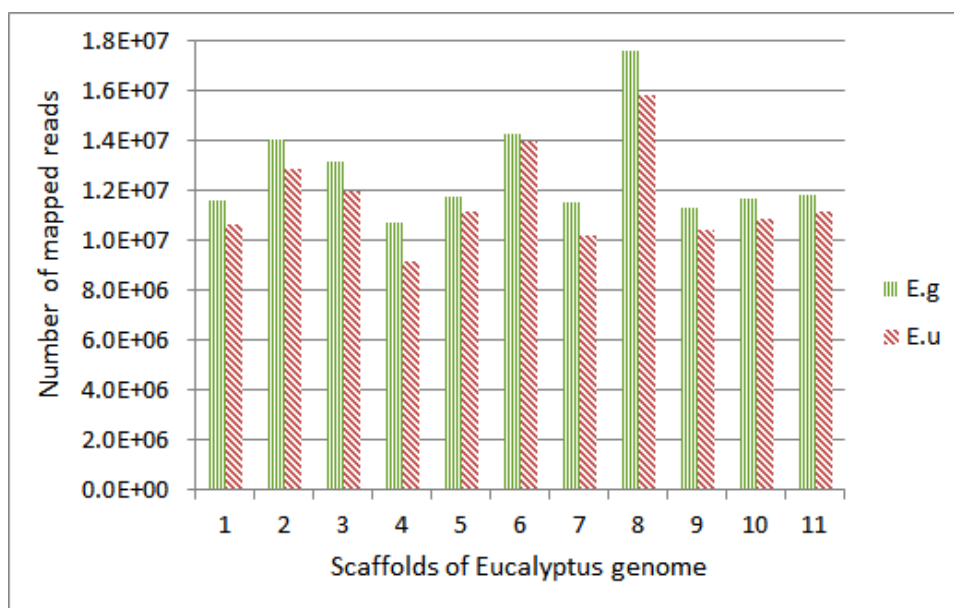


Figure S4.2 Distribution of mapped reads on the main 11 scaffolds of the BRASUZ1 genome sequence. Green bars for *E. grandis* (E.g) and red bars for *E. urophylla* (E.u).

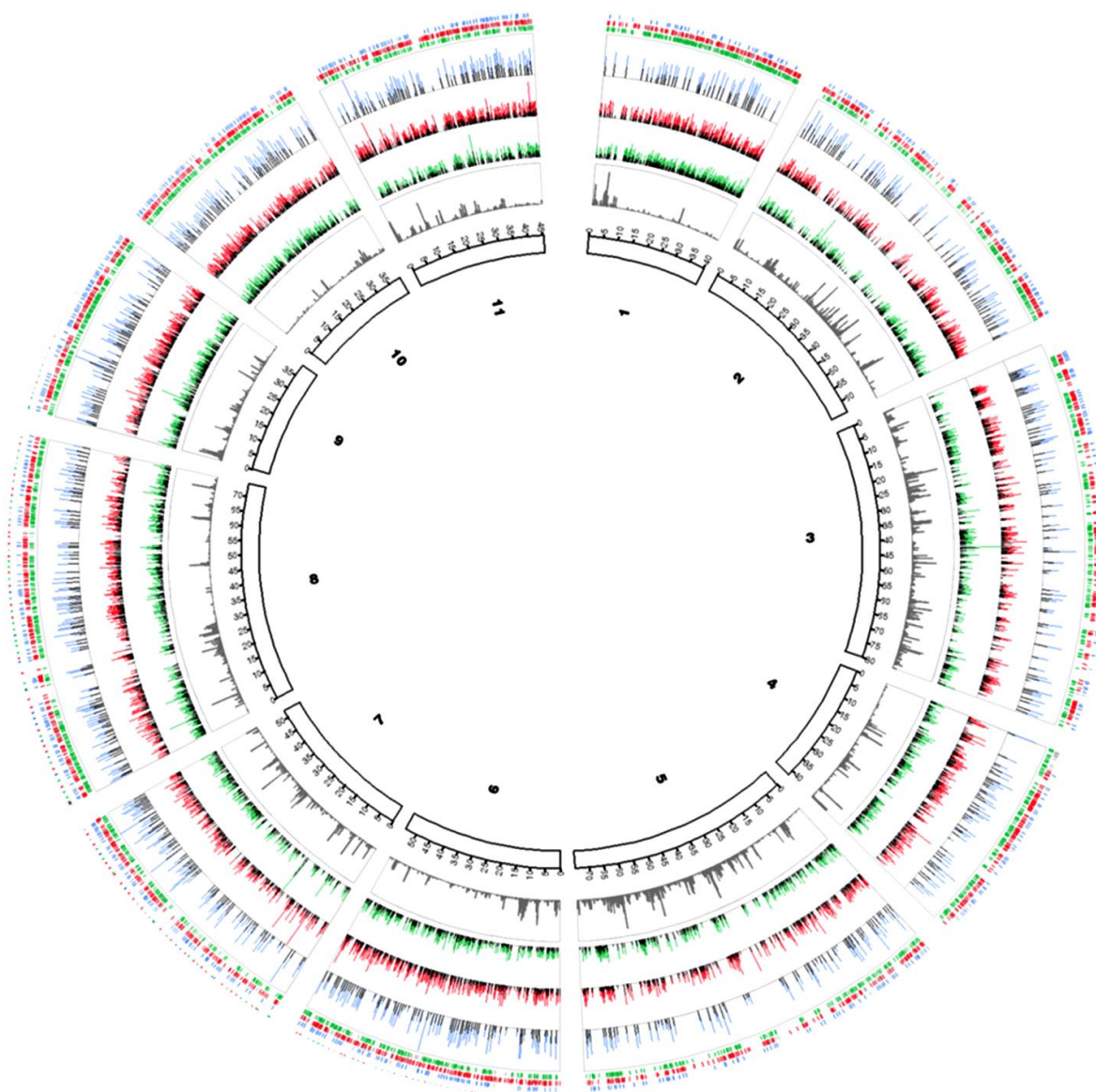


Figure S4.3 Distribution of selected SNPs along the main 11 scaffolds of the BRASUZ1 genome sequence. Distance in Mb (inner ring), percentage of missing values over a 200 kbp window (2nd ring), sequencing depth for each allele of set1 (3rd ring), set2 (4th ring) and set3 (5th ring), location of polymorphic SNPs (green: set1, red: set2 and blue: set3) and monomorphic SNPs (grey) (outer ring)

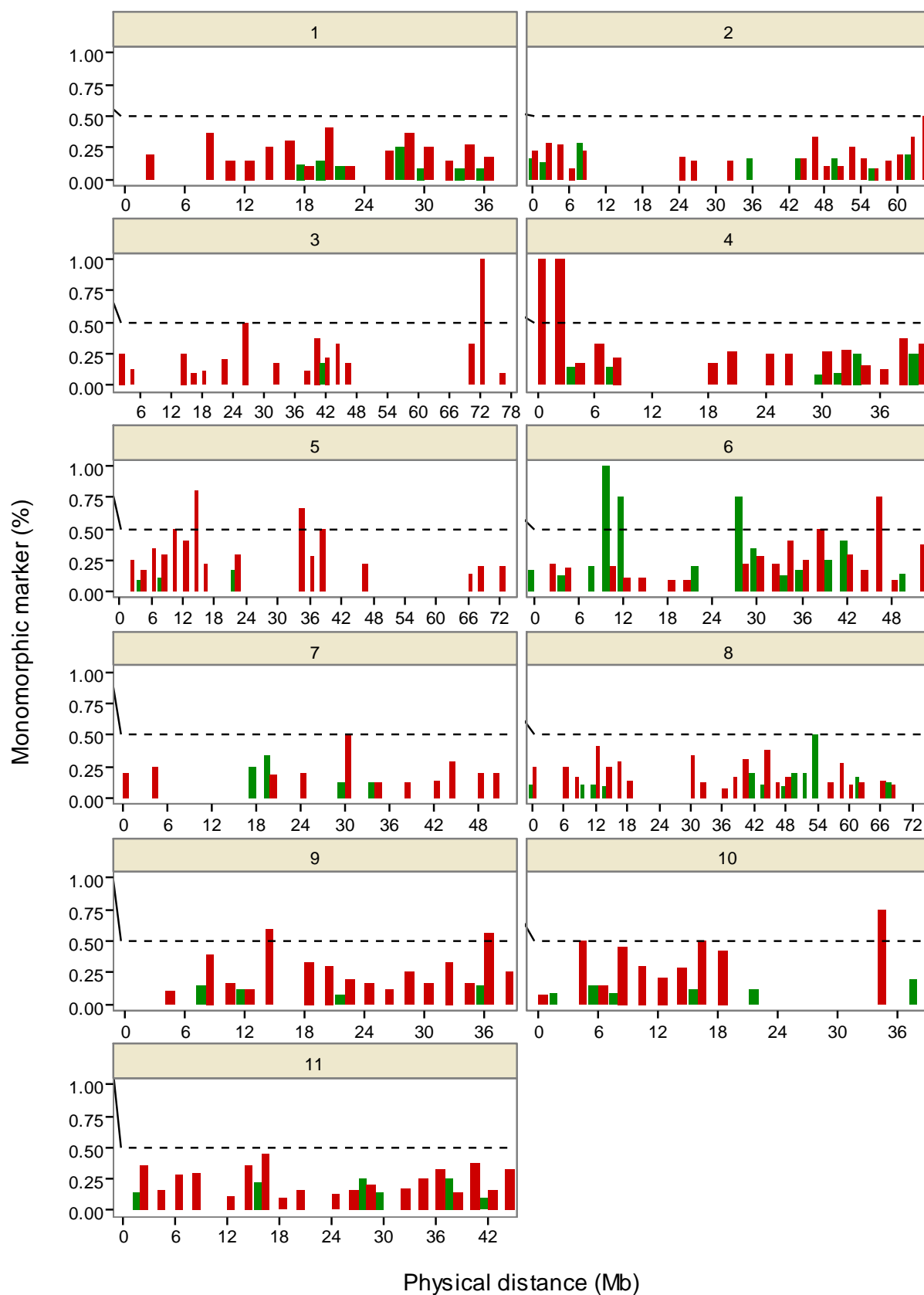


Figure S4.4 Distribution of the rate of monomorphic SNPs (per 2 Mb windows) along the main 11 scaffolds of the BRASUZ1 genome sequence *E. grandis* (E.g) in green and *E. urophylla* (E.u) in red.

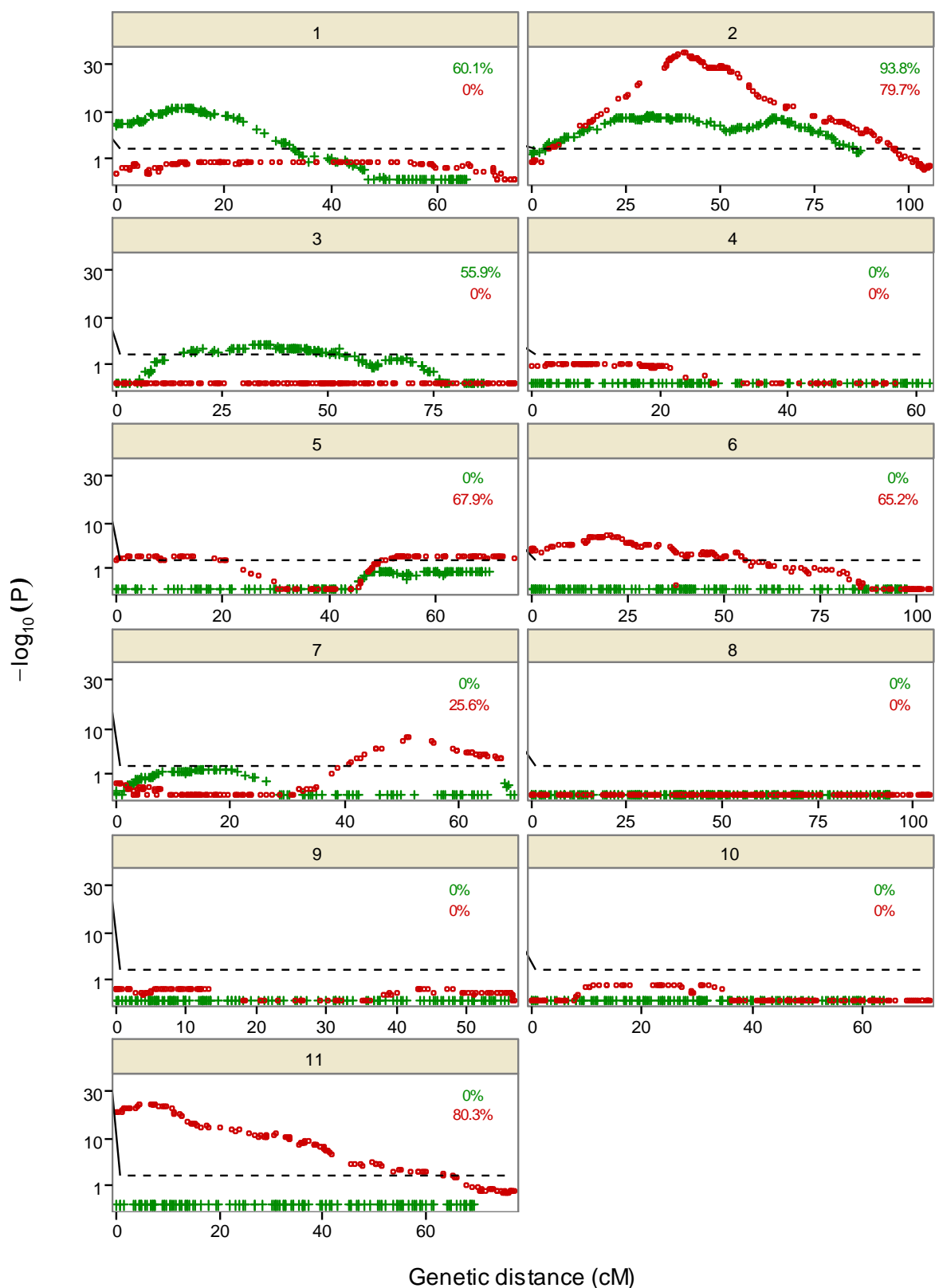


Figure S4.5 Evolution of SNP segregation along the 11 linkage groups of the Eucalyptus genome. Horizontal dashed lines represent the threshold above which SNPs are significantly distorted ( $P_{\text{adjust}} < 0.01$ ). *E. grandis* (*E.g*) is in green and *E. urophylla* (*E.u*) is in red.



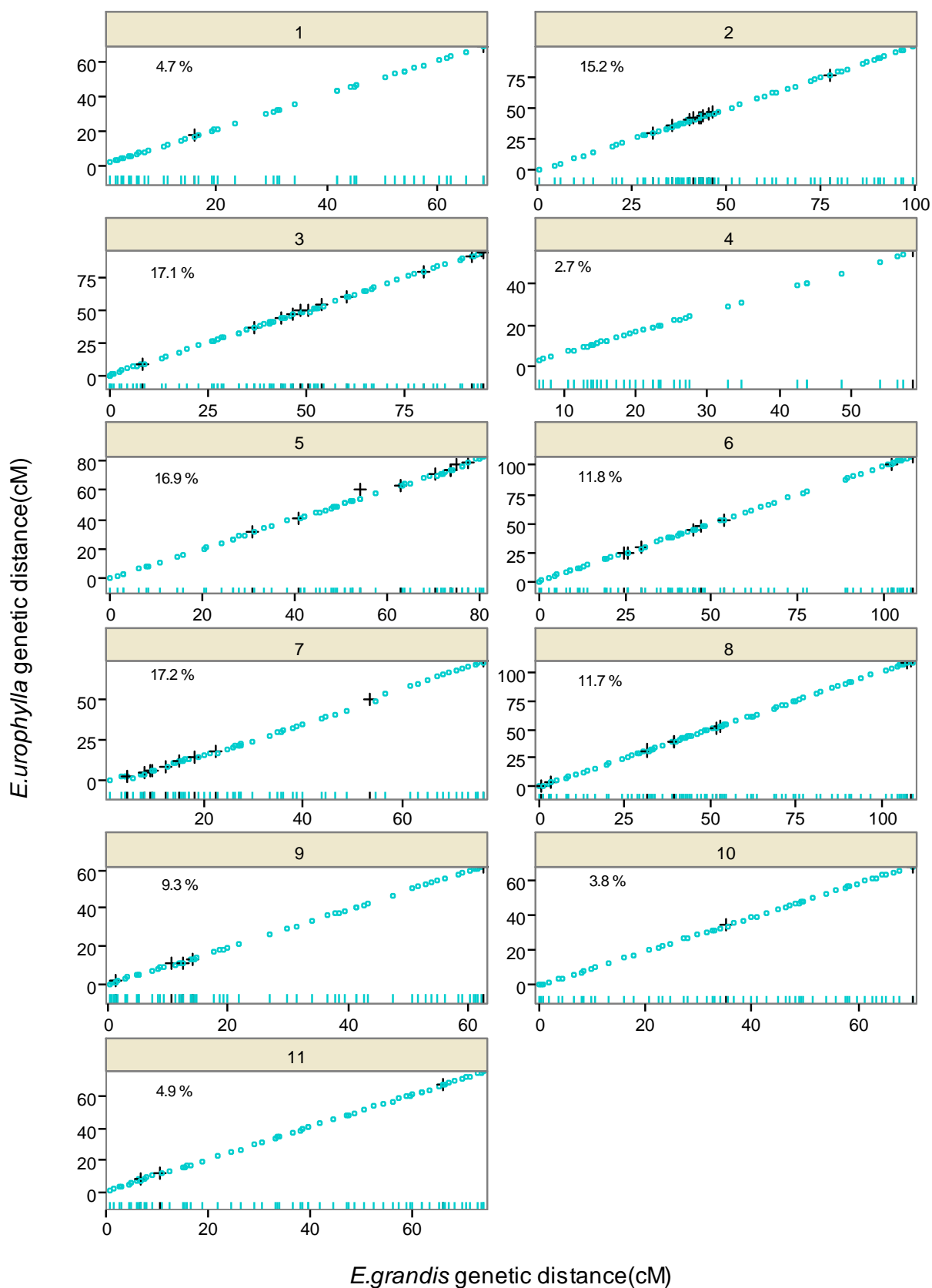


Figure S4.6 Comparison of the genetic location of SNPs segregating 1:2:1 (set3 in the material and methods section) between the two parental maps over the 11 linkage groups. Inverted SNPs between the two maps are indicated by black crosses.


Table S4.1 Comprehensive list of genetic linkage maps established for *Eucalyptus* species.

Species or hybrid	Pedigree	Number of markers	Type of markers	Number of LG	Map length (cM)	Type of map	Mapping software	References	Related reference
<i>E. grandis</i>	62 F1	240	RAPD	14	1,552	I	Mapmaker / Gmende1	(Grattapaglia & Sederoff, 1994)	
<i>E. urophylla</i>	62 F1	251	RAPD	11	1,101	I	Mapmaker / Gmende1	(Grattapaglia & Sederoff, 1994)	
<i>E. gunnii</i> x <i>E. globulus</i>	72 F2, 10 BC	97	RAPD	15	560*	C	Joinmap	(Vaillancourt <i>et al.</i> , 1995)	
<i>E. nitens</i>	118 F2	330	RFLP, RAPD, Isozyme	12	1,462	C	Joinmap / Crimap	(Byrne <i>et al.</i> , 1995)	
<i>E. urophylla</i>	93 F1	269	RAPD	11	1,331	I	Mapmaker / Joinmap	(Verhaegen & Plomion, 1996)	
<i>E. grandis</i>	93 F1	236	RAPD	11	1,415	I	Mapmaker / Joinmap	(Verhaegen & Plomion, 1996)	
<i>E. globulus</i>	91 F1	200	AFLP	16	967	I	PGRI / Mapmaker	(Marques <i>et al.</i> , 1998)	
<i>E. tereticornis</i>	91 F1	268	AFLP	14	919	I	PGRI / Mapmaker	(Marques <i>et al.</i> , 1998)	
<i>E. globulus</i> (female)	165 F1	120	RAPD, SSR	11	701	I	Mapmaker	(Bundock <i>et al.</i> , 2000)	
<i>E. globulus</i> (male)	165 F1	118	RAPD, SSR, Isozyme	13	1,013	I	Mapmaker	(Bundock <i>et al.</i> , 2000)	
<i>E. urophylla</i>	201 F1	8	Gene	5	-	I	Mapmaker	(Gion <i>et al.</i> , 2000)	Added to (Verhaegen & Plomion, 1996)
<i>E. grandis</i>	201 F1	4	Gene	3	-	I	Mapmaker	(Gion <i>et al.</i> , 2000)	
<i>E. globulus</i>	73 F1	+ 34	SSR	8	-	I	PGRI / Mapmaker	(Marques <i>et al.</i> , 2002)	Added to (Marques <i>et al.</i> , 1998)
<i>E. tereticornis</i>	73 F1	+ 34	SSR	8	-	I	PGRI / Mapmaker	(Marques <i>et al.</i> , 2002)	
<i>E. grandis</i>	92 F1	+ 63	SSR	11	-	I	Mapmaker	(Brondani <i>et al.</i> , 1998; Brondani <i>et al.</i> , 2002)	Added to (Grattapaglia & Sederoff, 1994)
<i>E. urophylla</i>	92 F1	+ 53	SSR	10	-	I	Mapmaker	(Brondani <i>et al.</i> , 1998; Brondani <i>et al.</i> , 2002)	
<i>E. camaldulensis</i>	92 F1	168	RAPD, RFLP, SSR	11	1,236	C	Joinmap	(Agrama <i>et al.</i> , 2002)	
<i>E. globulus</i>	148 F1	249	RFLP, EST, SSR, Gene, Isozyme	12	1,375	C	Joinmap	(Thamarus <i>et al.</i> , 2002)	
<i>E. urophylla</i>	82 F1	220	RAPD	23	1,505	I	Joinmap	(Gan <i>et al.</i> , 2003)	
<i>E. tereticornis</i>	82 F1	218	RAPD	23	1,036	I	Joinmap	(Gan <i>et al.</i> , 2003)	
<i>E. grandis</i>	156 BC	254	AFLP	11	1,335	I	Mapmaker	(Myburg <i>et al.</i> , 2003)	



Species or hybrid	Pedigree	Number of markers	Type of markers	Number of LG	Map length (cM)	Type of map	Mapping software	References	Related reference
<i>E. globulus</i>	177 BC	236	AFLP	11	1,405	I	Mapmaker	(Myburg <i>et al.</i> , 2003)	
<i>E. grandis</i> x <i>E. globulus</i> (male)	156 BC	342	AFLP	11	1,448	I	Mapmaker	(Myburg <i>et al.</i> , 2003)	
<i>E. grandis</i> x <i>E. globulus</i> (female)	177 BC	411	AFLP	11	1,318	I	Mapmaker	(Myburg <i>et al.</i> , 2003)	
<i>E. grandis</i>	92 F1	202	SSR	11	1,815	I	Joinmap	(Brondani <i>et al.</i> , 2006)	
<i>E. urophylla</i>	92 F1	160	SSR	11	1,333	I	Joinmap	(Brondani <i>et al.</i> , 2006)	
<i>E. grandis</i> x <i>E. urophylla</i>	92 F1	234	SSR	11	1,568	C	Joinmap	(Brondani <i>et al.</i> , 2006)	
<i>E. globulus</i> (female)	121 F2	118	AFLP, SSR	13	698	I	Joinmap	(Freeman <i>et al.</i> , 2006)	
<i>E. globulus</i> (male)	121 F2	130	AFLP, SSR	14	776	I	Joinmap	(Freeman <i>et al.</i> , 2006)	
<i>E. globulus</i>	121 F2	165	AFLP, SSR	11	702	C	Joinmap	(Freeman <i>et al.</i> , 2006)	
<i>E. nitens</i>	296 F2	296	RFLP, EST, SSR, Gene	11	1,700*	C	Outmap	(Henery <i>et al.</i> , 2007)	
<i>E. grandis</i> x <i>E. urophylla</i>	90 F1	51	RAPD	12	632	C	Gqmol	(Rocha <i>et al.</i> , 2007)	
<i>E. grandis</i> x <i>E. urophylla</i>	68 F1	2,053	SFP, SSR	11	1,330	C	Mapmaker / Joinmap	Neves 2011 (Neves <i>et al.</i> , 2011)	Added to (Brondani <i>et al.</i> , 2006)
<i>E. urophylla</i>	82 F1	268	EST-CAPS, RAPD	21	1,790	I	Mapmaker	Yu 2012 (Yu <i>et al.</i> , 2012)	Added to (Gan <i>et al.</i> , 2003)
<i>E. tereticornis</i>	82 F1	260	EST-CAPS, RAPD	21	1,488	I	Mapmaker	Yu 2012 (Yu <i>et al.</i> , 2012)	
<i>E. globulus</i> (GLOB-F1 -1)	184 F1	577	DArT, SSR, Gene	11	1,033	C	Joinmap	(Hudson <i>et al.</i> , 2012a; 2012b)	
<i>E. globulus</i> (GLOB-F1 -4)	184 F1	498	DArT, SSR, Gene	11	1,137	C	Joinmap	(Hudson <i>et al.</i> , 2012a; 2012b)	
<i>E. globulus</i> (GLOB-F1 -5)	183 F1	606	DArT, SSR, Gene	11	1,055	C	Joinmap	(Hudson <i>et al.</i> , 2012a; 2012b)	
<i>E. globulus</i> (GLOB-F2 -1)	172 F2	695	DArT, SSR, Gene	11	1,258	C	Joinmap	(Hudson <i>et al.</i> , 2012a; 2012b)	
<i>E. globulus</i> (GLOB-LH)	503 F2	1,060	DArT, SSR	11	1,151	C	Joinmap	(Hudson <i>et al.</i> , 2012a; 2012b)	
Reference map	-	4191	DArT, SSR, Gene	11	1,107	C	Joinmap	(Hudson <i>et al.</i> , 2012a; 2012b)	
<i>E. grandis</i>	180 BC	991	DArT, SSR	11	925	I	Joinmap	(Kullan <i>et al.</i> , 2012)	
<i>E. urophylla</i>	367 BC	958	DArT, SSR	11	1,107	I	Joinmap	(Kullan <i>et al.</i> , 2012)	
<i>E. grandis</i> x <i>E. urophylla</i> ( <i>E.g</i> BC)	180 BC	1,217	DArT, SSR	11	1,021	I	Joinmap	(Kullan <i>et al.</i> , 2012)	



Species or hybrid	Pedigree	Number of markers	Type of markers	Number of LG	Map length (cM)	Type of map	Mapping software	References	Related reference
<i>E. grandis</i> x <i>E. urophylla</i> (E.u BC)	367 BC	1,311	DArT, SSR	11	1,067	I	Joinmap	(Kullan <i>et al.</i> , 2012)	
<i>E. grandis</i> x <i>E. urophylla</i>	-	2,290	DArT, SSR	11	1,107	C	Joinmap	(Kullan <i>et al.</i> , 2012)	
<i>E. grandis</i> x <i>E. urophylla</i>	177 F1	2,484	DArT, SSR	11	1,304	C	Joinmap	(Petroli <i>et al.</i> , 2012)	
<i>E. grandis</i>	190 F1	86	RAPD	11	1,021	I	Joinmap	(Bartholomé <i>et al.</i> , 2013)	
<i>E. urophylla</i>	190 F1	106	RAPD	13	1,179	I	Joinmap	(Bartholomé <i>et al.</i> , 2013)	
<i>E. grandis</i>	1,025 F1	1,832	SNP	11	822	I	Onemap / Record	The present study	
<i>E. urophylla</i>	1,025 F1	1,773	SNP	11	886	I	Onemap / Record	The present study	

I :Individual map, C :consensus map

\* Estimate from figure. The authors did not provide the total map length (cM).

## References (Table S4.2)

- Agrama HA, George TL, Salah SF. 2002.** Construction of genome map for *Eucalyptus camaldulensis* DEHN. *Silvae genetica* **51**(5-6): 201-206.
- Bartholomé J, Salmon F, Vigneron P, Bouvet J-M, Plomion C, Gion J-M. 2013.** Plasticity of primary and secondary growth dynamics in *Eucalyptus* hybrids: a quantitative genetics and QTL mapping perspective. *BMC Plant Biology* **13**(1): 120.
- Brondani R, Brondani C, Tarchini R, Grattapaglia D. 1998.** Development, characterisation and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *Theoretical and Applied Genetics* **97**: 816 - 827.
- Brondani R, Williams E, Brondani C, Grattapaglia D. 2006.** A microsatellite-based C linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. *BMC Plant Biology* **6**(1): 20.
- Brondani RPV, Brondani C, Grattapaglia D. 2002.** Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. *Molecular Genetics and Genomics* **267**(3): 338-347.
- Bundock PC, Hayden M, Vaillancourt RE. 2000.** Linkage maps of *Eucalyptus globulus* using RAPD and microsatellite markers. *Silvae genetica* **49**(4-5): 223-232.
- Byrne M, Murrell JC, Allen B, Moran GF. 1995.** An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers. *Theoretical and Applied Genetics* **91**(6-7): 869-875.
- Freeman JS, Potts BM, Shepherd M, Vaillancourt RE. 2006.** Parental and C linkage maps of *Eucalyptus globulus* using AFLP and microsatellite markers. *Silvae genetica* **55**(4-5): 202-217.
- Gan S, Shi J, Li M, Wu K, Wu J, Bai J. 2003.** Moderate-Density Molecular Maps of *Eucalyptus Urophylla* S. T. Blake and *E. Tereticornis* Smith Genomes Based on RAPD Markers. *Genetica* **118**(1): 59-67.
- Gion J-M, Rech P, Grima-Pettenati J, Verhaegen D, Plomion C. 2000.** Mapping candidate genes in *Eucalyptus* with emphasis on lignification genes. *Molecular Breeding* **6**(5): 441-449.



- Grattapaglia D, Sederoff R. 1994.** Genetic Linkage Maps of *Eucalyptus grandis* and *Eucalyptus urophylla* Using a Pseudo-Testcross Mapping Strategy and RAPD Markers. *Genetics* **137**(4): 1121-1137.
- Henery ML, Moran GF, Wallis IR, Foley WJ. 2007.** Identification of quantitative trait loci influencing foliar concentrations of terpenes and formylated phloroglucinol compounds in *Eucalyptus nitens*. *New Phytologist* **176**(1): 82-95.
- Hudson CJ, Freeman JS, Kullan ARK, Petroli CD, Sansaloni CP, Kilian A, Detering F, Grattapaglia D, Potts BM, Myburg AA, et al. 2012a.** A reference linkage map for eucalyptus. *BMC Genomics* **13**(240): (15 June 2012)-(2015 June 2012).
- Hudson CJ, Kullan ARK, Freeman JS, Faria DA, Grattapaglia D, Kilian A, Myburg AA, Potts BM, Vaillancourt RE. 2012b.** High synteny and colinearity among *Eucalyptus* genomes revealed by high-density comparative genetic mapping. *Tree Genetics & Genomes* **8**(2): 339-352.
- Kullan ARK, van Dyk M, Jones N, Kanzler A, Bayley A, Myburg A. 2012.** High-density genetic linkage maps with over 2,400 sequence-anchored DArT markers for genetic dissection in an F2 pseudo-backcross of *Eucalyptus grandis* × *E. urophylla*. *Tree Genetics & Genomes* **8**(1): 163-175.
- Marques CM, Araújo JA, Ferreira JG, Whetten R, O'Malley DM, Liu BH, Sederoff R. 1998.** AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. *Theoretical and Applied Genetics* **96**(6-7): 727-737.
- Marques CM, Brondani RPV, Grattapaglia D, Sederoff R. 2002.** Conservation and synteny of SSR loci and QTLs for vegetative propagation in four *Eucalyptus* species. *Theoretical and Applied Genetics* **105**(2-3): 474-478.
- Myburg AA, Griffin AR, Sederoff RR, Whetten RW. 2003.** Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. *Theoretical and Applied Genetics* **107**(6): 1028-1042.
- Neves L, MC Mamani E, Alfenas A, Kirst M, Grattapaglia D. 2011.** A high-density transcript linkage map with 1,845 expressed genes positioned by microarray-based Single Feature Polymorphisms (SFP) in *Eucalyptus*. *BMC Genomics* **12**(1): 189.
- Petroli CD, Sansaloni CP, Carling J, Steane DA, Vaillancourt RE, Myburg AA, da Silva OB, Jr., Pappas GJ, Jr., Kilian A, Grattapaglia D. 2012.** Genomic Characterization of DArT Markers Based on High-Density Linkage Analysis and Physical Mapping to the *Eucalyptus* Genome. *PLoS ONE* **7**(9): e44684.
- Rocha RB, Barros EG, Cruz CD, Rosado AM, Araujo EFd. 2007.** Mapping of QTLs related with wood quality and developmental characteristics in hybrids ( *Eucalyptus grandis* \* *Eucalyptus urophylla*). *Revista Arvore* **31**(1): 13-24.
- Thamarus KA, Groom K, Murrell J, Byrne M, Moran GF. 2002.** A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre, and floral traits. *Theoretical and Applied Genetics* **104**(2-3): 379-387.
- Vaillancourt R, Potts B, Manson A, Eldridge T, Reid J 1995.** Using RAPDs to Detect QTLs in an Interspecific F2 Hybrid of *Eucalyptus*. In CRCTHF-IUFRO. *Eucalypt Plantations : Improving Fibre Yield and Quality - CRCTHF-IUFRO Conference*. Hobart. pp. 430-433.
- Verhaegen D, Plomion C. 1996.** Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers. *Genome* **39**(6): 1051-1061.
- Yu X, Guo Y, Zhang X, Li F, Weng Q, Li M, Gan S. 2012.** Integration of EST-CAPS markers into genetic maps of *Eucalyptus urophylla* and *E. tereticornis* and their alignment with *E. grandis* genome sequence. *Silvae genetica* **61**(6): 247-255.



Table S4.3 Detailed SNP genotyping results per scaffold and SNP set (set1: SNP segregating 1:1 in *E. grandis*; set2: SNPs segregating 1:1 in *E. urophylla*, set3: SNPs segregating 1:2:1 in both species)

SNP type	Scaffold	1	2	3	4	5	6	7	8	9	10	11	12	14	17	18	20	26	31	32	41	54	57	71	77	164	Total
<i>E.g</i> (1:1)	Polymorph	181	189	191	145	165	157	128	240	156	168	128	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1849
	Monomorph	9	9	1	7	3	19	4	12	4	7	7	-	-	-	-	-	-	-	-	-	0	-	-	-	-	82
	No call	12	20	12	6	22	13	17	18	17	9	11	-	-	-	-	-	-	-	-	-	0	-	-	-	-	157
	Total	202	218	204	158	190	189	149	270	177	184	146	-	-	-	-	-	-	-	-	-	1	-	-	-	-	2088
	% Polymorph	89.6	86.7	93.6	91.8	86.8	83.1	85.9	88.9	88.1	91.3	87.7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	88.6
	% Monomorph	4.5	4.1	0.5	4.4	1.6	10.1	2.7	4.4	2.3	3.8	4.8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.9
<i>E.u</i> (1:1)	Polymorph	141	206	182	125	157	175	151	226	124	145	161	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1793
	Monomorph	35	34	23	35	34	38	18	37	38	31	43	-	-	-	-	-	-	-	-	-	-	-	-	-	-	366
	No call	11	15	11	10	17	6	18	18	8	16	9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	139
	Total	187	255	216	170	208	219	187	281	170	192	213	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2298
	% Polymorph	75.4	80.8	84.3	73.5	75.5	79.9	80.7	80.4	72.9	75.5	75.6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	78
	% Monomorph	18.7	13.3	10.6	20.6	16.3	17.4	9.6	13.2	22.4	16.1	20.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	15.9
1:2:1	Polymorph	41	76	79	38	69	72	69	88	58	56	66	1	1	2	1	1	1	1	1	1	-	1	1	0	1	725
	Monomorph	0	0	2	3	2	5	1	3	1	3	1	0	0	0	0	0	0	0	0	0	-	0	0	0	0	21
	No call	3	4	2	4	6	6	5	7	3	2	0	0	0	0	0	0	0	0	0	0	-	0	0	1	0	43
	Total	44	80	83	45	77	83	75	98	62	61	67	1	1	2	1	1	1	1	1	1	-	1	1	1	1	789
	% Polymorph	93.2	95	95.2	84.4	89.6	86.7	92	89.8	93.5	91.8	98.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	91.9
	% Monomorph	0	0	2.4	6.7	2.6	6	1.3	3.1	1.6	4.9	1.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.7
Total	% No Call	6.8	5	2.4	8.9	7.8	7.2	6.7	7.1	4.8	3.3	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5.4
	Total	433	553	503	373	475	491	411	649	409	437	426	1	1	2	1	1	1	1	1	1	1	1	1	1	1	5175

Table S4.4 Location of monomorphic regions on the *Eucalyptus* genome and the genotyping results of the associated SNPs. Polymorphic (P), monomorphic (M) and no call (NC)

Parent	LG	Location of monomorphic marker		Dist (Mb)	<i>E.g</i>			<i>E.u</i>			1:2:1		
		First	Last		P	M	NC	P	M	NC	P	M	NC
<i>E.g</i>	6	9809041	13701289	3.89	0	5	1	9	2	0	0	1	2
	6	28960384	30414900	1.45	0	5	0	5	3	0	0	0	2
	6	41704192	43408720	1.7	0	3	1	5	2	0	0	1	0
<i>E.u</i>	4	52127	3829866	3.78	16	0	0	0	11	0	0	0	0
	5	14713336	15784895	1.07	5	0	0	0	5	1	0	2	0
	6	46144072	47524696	1.38	9	0	0	0	5	1	0	0	0
	10	16661541	18761220	2.1	8	1	0	0	5	0	0	1	0
	10	32621512	35186096	2.56	2	0	0	0	2	5	0	0	0



Table S4.5 Framework linkage map features based on test-cross markers only (1:1 segregation) for *E. grandis* (*E.g*) and *E. urophylla* (*E.u*). Two algorithms ML (JoinMap) and RECORD (Onemap) are compared.

	LG	Scaffold											Total
		1	2	3	4	5	6	7	8	9	10	11	
<i>E.g</i>	ML	40,297,282	64,237,462	80,088,348	41,978,404	74,731,017	53,893,726	52,447,651	74,330,457	39,019,482	39,359,118	45,510,589	605,893,536
	Total map length (cM)	65.27	87.37	87.14	61.8	69.78	97.08	69.78	93.81	56.06	63.6	69.44	821.13
	Number of SNPs	183	176	195	136	169	164	134	230	156	162	127	1832
	Number of bins	132	151	151	100	130	129	120	174	115	126	104	1432
	Distance between SNPs (cM)	0.357	0.496	0.447	0.454	0.413	0.592	0.521	0.408	0.359	0.393	0.547	0.453
	Distance between bins (cM)	0.494	0.579	0.577	0.618	0.537	0.753	0.582	0.539	0.487	0.505	0.668	0.576
	Distorted SNPs (%)	60.1	93.8	55.9	0	0	0	0	0	0	0	0	21
	Coverage (%)	99.7	97.4	95.6	97.6	99.2	98.7	98.7	99.3	99.5	99.2	96.1	98.2
	RECORD Total map length (cM)	65.25	87.45	87.03	61.9	70.23	97.32	69.74	93.95	55.9	63.59	69.3	821.66
	Number of SNPs	183	176	195	136	169	164	134	230	156	162	127	1832
	Number of bins	132	151	151	101	131	131	115	174	114	125	104	1429
<i>E.u</i>	ML	74.05	105.36	94.44	56.29	74.98	103.48	67.26	103.55	56.84	71.91	76.74	884.9
	Total map length (cM)	146	192	186	117	159	181	156	219	121	139	157	1773
	Number of SNPs	105	149	147	87	113	141	110	160	95	118	119	1344
	Number of bins	105	149	147	87	113	141	110	160	95	118	119	1344
	Distance between SNPs (cM)	0.507	0.549	0.508	0.481	0.472	0.572	0.431	0.473	0.470	0.517	0.489	0.497
	Distance between bins (cM)	0.705	0.707	0.642	0.647	0.664	0.734	0.611	0.647	0.598	0.609	0.645	0.655
	Distorted SNPs (%)	0	79.7	0	0	67.9	65.2	25.6	0	0	0	80.3	30.7
	Coverage (%)	99.1	97.2	94.4	85.6	99	99.6	98.8	99.8	97.2	99.3	97.2	97.2
	RECORD Total map length (cM)	74.19	105.45	94.28	56.51	74.81	103.27	67.45	104.14	56.81	71.86	77.15	885.92
	Number of SNPs	146	192	186	117	159	181	156	219	121	139	157	1773
	Number of bins	107	149	148	89	113	141	113	161	95	118	119	1353
	Distance between SNPs (cM)	0.508	0.549	0.507	0.483	0.471	0.571	0.432	0.476	0.470	0.517	0.491	0.498
	Distance between bins (cM)	0.693	0.708	0.637	0.635	0.662	0.732	0.597	0.647	0.598	0.609	0.648	0.652
	Distorted SNPs (%)	0	79.7	0	0	67.9	65.2	25.6	0	0	0	80.3	30.7
	Coverage (%)	99.1	97.2	94.4	85.6	99	99.6	98.8	99.8	97.2	99.3	97.2	97.2





Table S4.6 Linkage map features based on test-cross (1:1 segregation) and inter-cross (1:2:1) markers for *E. grandis* (E.g) and *E. urophylla* (E.u). The RECORD algorithm was used.

LG	1	2	3	4	5	6	7	8	9	10	11	Total
Scaffold size (bp)	40,297,282	64,237,462	80,088,348	41,978,404	74,731,017	53,893,726	52,447,651	74,330,457	39,019,482	39,359,118	45,510,589	605,893,536
<i>E. g</i>												
Map length (cM)	70.72	99.95	95.14	62.46	81.67	108.22	76.29	109.07	63.73	70.18	75.16	912.59
Number of SNPs (total)	228	252	277	174	238	240	203	316	216	217	191	2,551
Number of SNPs (1:1)	183	176	195	136	169	164	134	230	156	162	127	1,832
Number of SNPs (1:2:1)	45	76	82	38	69	76	68	86	60	55	64	719
Number of bins	168	201	199	141	173	183	162	233	148	155	141	1,904
Distance between SNP (cM)	0.31	0.40	0.34	0.36	0.34	0.45	0.38	0.35	0.30	0.32	0.39	0.36
Distance between bin (cM)	0.42	0.50	0.48	0.44	0.47	0.59	0.47	0.47	0.43	0.45	0.53	0.48
Distorted SNPs (%)	59.2	85.7	37.5	0	14.3	0	0	0	0	0	20.9	19.8
Coverage (%)	99.7	97.4	95.7	97.6	99.3	98.7	98.8	99.5	99.5	99.2	96.1	98.3
<i>E. u</i>												
Map length (cM)	71.31	99.78	95.34	57.13	85.1	107.51	72.53	109.06	62.03	68.31	75.89	903.99
Number of SNPs (total)	191	269	268	155	228	257	224	305	181	194	221	2,491
Number of SNPs (1:1)	146	192	186	117	159	181	156	219	121	139	157	1,773
Number of SNPs (1:2:1)	45	76	82	38	69	76	68	86	59	55	64	718
Number of bins	134	199	188	118	158	192	146	214	134	146	170	1,799
Distance between SNP (cM)	0.37	0.37	0.36	0.37	0.37	0.42	0.32	0.36	0.34	0.35	0.34	0.36
Distance between bin (cM)	0.53	0.5	0.51	0.48	0.54	0.56	0.5	0.51	0.46	0.47	0.45	0.5
Distorted SNPs (%)	10.5	75.5	0	0	62.7	52.1	21	0	0	0	76.5	28.7
Coverage (%)	99.1	97.2	95.6	85.6	99	99.6	98.8	99.9	99.1	99.4	97.2	97.3



Table S4.7 Characteristics of the four marker density classes (MD) based on framework maps for *E. grandis* (E.g) and *E. urophylla* (E.u).

		<i>E.g</i>				<i>E.u</i>			
		MD1	MD2	MD3	MD4	MD1	MD2	MD3	MD4
Nb of marker		175	337	734	1,832	187	352	734	1,773
Marker (%)		9.6	18.4	40.1	100	10.5	19.9	41.4	100
Distance between bin (cM)	- mean	5.01	2.52	1.14	0.58	5.05	2.6	1.23	0.66
	- sd	0.61	0.58	0.51	0.56	0.77	0.74	0.71	0.72
	- maximum	7.4	4.41	4.28	4.28	7.18	6.24	6.24	6.24
	- minimum	3.25	1.19	0.3	0.01	2.57	1.08	0.3	0.01
Missing data (%)		0.01	0.01	0.01	< 0.01	0.02	0.02	0.01	< 0.01

Table S4.8 Averages and standard deviations for 1,000 genetic maps established for different modalities of marker density (MD) and sample size (100, 200, 500).

Sample size		<i>E.g</i>			<i>E.u</i>		
Marker density	Sample size	Map length (cM)	Number of bin	Distance between bin (cM)	Map length (cM)	Number of bin	Distance between bin (cM)
MD1	100	851 (25)	173.9 (1.1)	4.9 (0.14)	921.8 (27.1)	185.8 (1.1)	5 (0.15)
MD1	200	846.2 (15.7)	175 (0.1)	4.8 (0.09)	917.1 (16.7)	187 (0.1)	4.9 (0.09)
MD1	500	844 (7.8)	175 (0)	4.8 (0.05)	913.9 (8.7)	187 (0)	4.9 (0.05)
MD2	100	834.5 (24.5)	309.1 (5)	2.7 (0.07)	900.6 (25.7)	322.8 (4.9)	2.8 (0.08)
MD2	200	830 (15.5)	334.5 (1.6)	2.5 (0.05)	896.4 (15.9)	349.1 (1.7)	2.6 (0.05)
MD2	500	828.1 (7.7)	337 (0)	2.5 (0.02)	893.4 (8.4)	352 (0.1)	2.5 (0.02)
MD3	100	826.8 (24.1)	487 (11.7)	1.7 (0.04)	892.8 (25.2)	493.5 (11.4)	1.8 (0.04)
MD3	200	822.7 (15.4)	644.2 (8.1)	1.3 (0.02)	888.7 (15.8)	644.1 (7.9)	1.4 (0.02)
MD3	500	820.8 (7.7)	729.8 (2.1)	1.1 (0.01)	885.8 (8.3)	728.7 (2.1)	1.2 (0.01)
MD4	100	827.1 (25)	567.9 (17.1)	1.5 (0.03)	893.1 (25.1)	560.6 (14.5)	1.6 (0.04)
MD4	200	823.4 (16.6)	844 (17.5)	1 (0.02)	889.2 (15.7)	809.1 (14.2)	1.1 (0.02)
MD4	500	822.7 (8.8)	1212.5 (14.5)	0.7 (0.01)	886.5 (8.3)	1135.5 (11.4)	0.8 (0.01)
Framework map		821.66	1429	0.578	885.92	1353	0.652



Table S4.9 Physical coordinates of non-colinear regions (NCR) on the BRASUZ1 genome sequence.

Sca	Map	S	N	Marker Index		Position (bp)		Position (cM)		Type of non-colinearity	Observation
				first	last	first	last	first	last		
1	<i>E.g</i>	3	3	3074	3226	13023	371529	0.88	1.08	inverted segment	
1	<i>E.u</i>	3	3	30	124	16093	326129	1.87	2.56		
1	<i>E.g</i>	2	3	3261	3073	879689	1298668	0.1	0.39	displaced segment	
1	<i>E.u</i>	2	2	8	17	1168498	1341934	0.99	1.78		
1	<i>E.u</i>	1	2	76	103	2444644	2930365	0	0	displaced segment	no <i>E.g</i> marker in the region
1	<i>E.g</i>	1	3	3171	3204	2943812	3415027	0	0.1	displaced segment	no <i>E.u</i> marker in the region
1	<i>E.g</i>	4	3	3249	3251	4923675	5261004	2.66	2.85	inverted segment	no <i>E.u</i> marker in the region
1	<i>E.u</i>	4	5	169	173	5263533	6259952	3.25	3.84	inverted segment	no <i>E.g</i> marker in the region
1	<i>E.g</i>	6	34	3255	3088	7936202	15268541	4.13	11.31	inverted segment	
1	<i>E.u</i>	6	23	175	25	8070538	15252341	5.93	13.41		
1	<i>E.g</i>	9	14	3091	3104	15535547	17342690	29.84	34.76	displaced segment	
1	<i>E.u</i>	9	8	27	38	15783048	17437207	37.12	41.43		
1	<i>E.g</i>	7	14	2332	3122	17474755	20970572	13.67	15.84	inverted segment	
1	<i>E.u</i>	7	15	39	59	17910828	20949351	19.25	23		
1	<i>E.g</i>	5	4	3123	3126	21180525	22264501	3.94	4.13	displaced segment	
1	<i>E.u</i>	5	2	60	64	21236460	22149053	5.73	5.83		
1	<i>E.g</i>	8	27	3127	3157	22556501	26743668	15.84	27.44	inverted segment	
1	<i>E.u</i>	8	19	65	89	22357126	26552765	23.69	34.72		
2	<i>E.g</i>	1	21	3636	3660	22624535	27166935	30.03	32.49	inverted segment	
2	<i>E.u</i>	1	24	633	663	22607963	27263099	47.49	49.55		
2	<i>E.u</i>	3	2	666	667	28755000	29800946	49.75	50.14	inverted segment	Only one <i>E.g</i> marker in the region
2	<i>E.g</i>	4	3	668	3669	31281098	32176641	36.52	36.91	inverted segment	
2	<i>E.u</i>	5	5	669	674	31830967	32486380	51.22	51.71		
2	<i>E.g</i>	2	4	3671	3675	32764035	33401768	32.49	33.37	inverted segment	linked markers
2	<i>E.u</i>	2	3	675	678	32789998	33363987	49.65	49.65	-	
2	<i>E.g</i>	3	2	3689	3690	40317577	40458277	33.96	34.06	inverted segment	
2	<i>E.u</i>	4	2	693	694	40338707	40457792	50.14	50.43		
2	<i>E.g</i>	5	2	3724	3726	49316002	49513840	58.84	60.02	inverted segment	
2	<i>E.u</i>	6	4	732	736	49340931	49679428	81.59	82.37		
2	<i>E.g</i>	6	8	3741	3747	52248929	53276680	60.22	63.68	inverted segment	
2	<i>E.u</i>	7	6	750	758	52254518	53369841	82.57	85.13		
2	<i>E.g</i>	7	12	3763	3775	57333078	59200036	68.3	74.71	inverted segment	
2	<i>E.u</i>	8	12	777	789	57339737	59093185	92.43	97.78		



3	<i>E.g</i>	7	5	3853	3857	19070366	19660445	47.85	48.63	inverted	segment	
3	<i>E.u</i>	5	5	890	895	18752518	19655164	51.41	52.8			
3	<i>E.g</i>	1	2	3858	3859	19904656	20075076	29.02	30.11	inverted	segment	Only one <i>E.u</i> marker in the region
3	<i>E.g</i>	2	4	3861	3864	22044378	23443882	35.36	36.34			
3	<i>E.u</i>	1	6	898	904	21984593	24634677	35.17	37.53	inverted	segment	
3	<i>E.g</i>	3	2	905	3867	25576935	26140670	37.92	38.22			
3	<i>E.g</i>	4	16	3868	3882	27938323	30927719	38.51	41.76	inverted	segment	
3	<i>E.u</i>	2	9	911	922	28607414	30926239	40.6	43.65			
3	<i>E.g</i>	5	6	3883	3891	32183335	34553062	41.76	42.25	inverted	segment	
3	<i>E.u</i>	3	5	924	930	32525738	33839729	43.94	44.73			
3	<i>E.g</i>	6	5	3924	3930	42644827	44347921	45.78	47.26	inverted	segment	
3	<i>E.u</i>	4	8	962	972	42559564	44258105	49.26	50.83			
3	<i>E.g</i>	8	5	1015	3979	61018597	62506118	61.76	68.78	inverted	segment	
3	<i>E.u</i>	6	3	1017	1019	61948294	62908685	75.71	76.8			
3	<i>E.u</i>	7	2	1025	1027	66593744	67490521	77.09	77.09	displaced	segment	Only one <i>E.g</i> marker in the region  no <i>E.u</i> marker in the region
3	<i>E.g</i>	9	3	3998	4000	72683040	73347497	72.64	74.02			
3	<i>E.g</i>	1 0	6	4002	4009	74459707	75518125	84.76	87.03	inverted	segment	
3	<i>E.u</i>	8	11	1047	1058	74487565	75786807	93.01	94.28			
4	<i>E.g</i>	1	6	4047	4053	15381742	16713580	15.89	16.77	inverted	segment	
4	<i>E.u</i>	1	7	1104	1111	15136006	16587761	8.66	9.35			
4	<i>E.g</i>	2	8	4084	4093	23432455	25506808	17.36	21.8	displaced	segment	
4	<i>E.u</i>	2	6	1146	1153	23547220	25445631	10.04	14.19			
5	<i>E.g</i>	1	7	4218	4224	18703374	20421339	39.03	40.71	inverted	segment	
5	<i>E.u</i>	1	8	1284	1294	18645503	20348533	36.84	40.39			
5	<i>E.g</i>	2	4	4244	4247	28735985	29417605	46.42	46.91	inverted	segment	
5	<i>E.u</i>	2	5	1315	1320	28545377	29421927	47.36	47.95			
5	<i>E.g</i>	3	12	4249	4261	34928546	38075742	48.58	50.65	inverted	segment	
5	<i>E.u</i>	3	10	1328	1343	35915942	38043813	51.69	52.77			
5	<i>E.g</i>	4	18	4268	4293	42579686	47975672	50.75	54.78	inverted	segment	
5	<i>E.u</i>	4	22	1349	1376	41431024	47668180	53.36	56.5			
5	<i>E.g</i>	5	4	4330	4334	62619948	65590631	61.58	62.07	inverted	segment	
5	<i>E.u</i>	5	8	1416	1423	62593790	66960182	66.42	67.31			
5	<i>E.g</i>	6	17	4335	4354	67169796	70305090	62.47	68.66	inverted	segment	
5	<i>E.u</i>	6	15	1424	1446	67158895	72073616	67.31	72.13			
6	<i>E.g</i>	1	3	4557	4568	545538	841908	0	0.39	inverted	segment	
6	<i>E.u</i>	1	2	1663	1675	545682	800616	0	0.1			
6	<i>E.g</i>	2	10	4397	4472	1750224	3478202	0.39	3.64	inverted	segment	
6	<i>E.u</i>	2	9	1489	1577	1728501	3712164	0.1	3.96			



6	<i>E.g</i>	3	7	4508	4558	4442727	5694831	4.72	8.28	inverted segment	
6	<i>E.u</i>	3	9	1591	1666	4007944	5906828	4.15	10.19		
6	<i>E.g</i>	4	2	4559	4560	6053542	6213364	8.48	9.37	inverted segment	
6	<i>E.u</i>	4	3	1667	1669	6066215	6274519	10.19	10.78		
6	<i>E.u</i>	5	3	1485	1487	16431281	17061960	25.51	25.61	inverted segment	Only one <i>E.g</i> marker in the region
6	<i>E.g</i>	5	9	4423	4431	24067212	25337256	40.05	40.15	displaced segment	
6	<i>E.u</i>	6	5	1517	1521	23913579	25346592	44.75	44.75	-	linked markers
6	<i>E.g</i>	6	12	4453	4468	30772049	33220694	40.15	42.81	inverted segment	
6	<i>E.u</i>	7	18	1536	1562	28817930	33211559	45.44	53.25		
7	<i>E.g</i>	1	2	4610	2874	253093	394315	0	0.1	inverted segment	no <i>E.u</i> marker in the region
7	<i>E.u</i>	1	4	1860	1865	7847575	8600813	15.85	17.13	inverted segment	Only one <i>E.g</i> marker in the region
7	<i>E.u</i>	2	4	1820	4688	45926379	46510154	59.65	61.84	inverted segment	Only one <i>E.g</i> marker in the region
8	<i>E.g</i>	1	16	4814	4946	3526444	6606892	0	3.44	displaced segment	
8	<i>E.u</i>	1	10	1962	2080	3579890	6018053	0	2.17		
8	<i>E.g</i>	2	13	4755	4771	15017801	19238010	27.53	34.76	inverted segment	
8	<i>E.u</i>	2	12	1891	1911	15075636	19331309	27.07	36.72		
8	<i>E.g</i>	3	2	4774	4775	20283800	20619378	37.32	37.91	displaced segment	
8	<i>E.u</i>	3	3	1913	1918	20366633	21952652	40.47	41.55		
8	<i>E.g</i>	4	7	4871	4878	48033095	49481233	75.87	80.32	displaced segment	
8	<i>E.u</i>	4	6	2024	2034	48140106	49488666	90.09	93.66		
8	<i>E.g</i>	5	8	4977	4984	72468970	74314569	93.07	93.95	inverted segment	
8	<i>E.u</i>	5	5	2130	2136	72321656	74272147	103.65	104.14		
10	<i>E.g</i>	1	7	3287	3293	13347150	14372161	30.07	30.46	inverted segment	
10	<i>E.u</i>	1	5	211	221	13291249	14380318	36.6	37.49		
11	<i>E.g</i>	1	5	3482	3488	19030955	20560968	26.97	30.84	inverted segment	
11	<i>E.u</i>	1	9	428	439	18953145	20608325	29.26	33.41		
11	<i>E.u</i>	2	2	459	460	26169849	26275838	33.7	33.7	displaced segment	Only one <i>E.g</i> marker in the region
11	<i>E.g</i>	2	3	3530	3532	32795211	33971650	34.38	35.57	inverted segment	
11	<i>E.u</i>	3	6	499	506	32795891	33870317	38.35	38.84		

scaffold: Sca, segment id: S, number of marker: N



Table S4.10 Physical coordinates of the non-syntenic regions (NSR) on the BRASUZ1 genome sequence.

Sca	LG	Map	Nb	Marker index		Position (bp)		Position (cM)		Final orientation	Observation
				first	last	first	last	first	last		
1	7	GRA	1	3089	3089	1528024	1528024	4.54	4.54	-	no close <i>E.u</i> marker in the region
1	6	GRA	3	3252	3254	7375679	7746423	39.92	39.92	-	
1	6	URO	1	174	174	7578707	7578707	45.36	45.36	-	
1	3	GRA	1	3158	3158	26902675	26902675	58.29	58.29	-	
1	3	URO	1	91	91	26862418	26862418	59.99	59.99	-	
2	3	URO	2	595	664	103305	274967	54.34	54.34	-	no close <i>E.g</i> marker in the region - scaffold start
2	6	GRA	1	3812	3812	824533	824533	45.11	45.11	-	
2	6	URO	3	825	600	640838	1221624	53.41	54.4	inverted	
2	1	URO	1	692	692	38930748	38930748	11.69	11.69	-	no close <i>E.g</i> marker in the region
2	1	GRA	8	3727	3739	49754256	51589806	11.37	13.64	inverted	
2	1	URO	10	737	749	49774632	51430825	15.48	19.24		
2	3	GRA	1	3749	3749	54190153	54190153	52.91	52.91	-	no close <i>E.u</i> marker in the region
4	5	GRA	2	4039	4040	13349631	13723285	68.2	68.2	-	no close <i>E.u</i> marker in the region
4	5	GRA	5	4101	4105	29729507	30273868	47.34	47.73	identical	
4	5	URO	4	1164	1167	29729334	30069832	48.08	48.87		
4	2	URO	1	1190	1190	34580515	34580515	47.71	47.71	-	
4	2	URO	1	1221	1221	41293213	41293213	43.85	43.85	-	no close <i>E.g</i> marker in the region - scaffold end
5	6	GRA	2	1250	4183	10315656	10633525	94.12	94.51	identical	
5	6	URO	1	1251	1251	10574859	10574859	101.29	101.29	-	
5	9	URO	1	1379	1379	48764083	48764083	17.43	17.43	-	no close <i>E.g</i> marker in the region
5	3	GRA	1	4328	4328	60450736	60450736	36.61	36.61	-	
5	3	URO	2	1410	1412	60068005	60299719	38.31	38.41	inverted	
6	3	GRA	1	4469	4469	34003205	34003205	53.51	53.51	-	close <i>E.u</i> marker well positioned
7	11	GRA	1	4589	4589	17257636	17257636	6.94	6.94	-	
7	3	GRA	1	4714	4714	52222610	52222610	42.25	42.25	-	close <i>E.u</i> marker well positioned
8	6	GRA	3	4734	4738	11263786	11816714	8.62	10.5	identical	close <i>E.u</i> marker well positioned
8	1	GRA	1	4778	4778	23849586	23849586	2.28	2.28	-	



8	1	URO	1	1920	1920	23883262	23883262	3.17	3.17	-	
8	7	GRA	1	4870	4870	47543388	47543388	10.78	10.78	-	
8	7	URO	1	2023	2023	47560157	47560157	13.75	13.75	-	
8	7	GRA	5	4971	4976	70998055	71743127	19.07	20.16	identical	
8	7	URO	2	2126	2127	71112510	71281010	23.74	23.74	-	
9	2	URO	1	2319	2319	9118129	9118129	50.07	50.07	-	no close <i>E.g</i> marker in the region
9	6	URO	1	2233	2233	26948161	26948161	37.89	37.89	-	close <i>E.u</i> markers well positioned - close <i>E.g</i> marker displaced in the same scaffold
10	8	GRA	2	3294	3296	15362002	15686400	92.63	92.82	identical	
10	8	URO	1	225	225	15697116	15697116	103.06	103.06	-	
10	7	GRA	3	3385	3388	32372310	32715204	12.16	12.25	identical	
10	7	URO	2	309	310	32314937	32470045	3.16	14.54		
11	5	GRA	2	3501	3502	24828540	25299926	48.13	48.23		
11	5	URO	4	452	456	24700582	25445386	49.07	50.55	inverted	

Table S4.11 Map location of additional scaffolds. The physical location of the flanking SNP markers is also provided.

Sca	Sca size	Marker id	Position (bp)	LG	Position(cM)		Flanking marker position			
					<i>E.g</i>	<i>E.u</i>	<i>E.g</i>		<i>E.u</i>	
							cM	bp	cM	bp
12	1,427,994	2490	192697	1	1.92	2.81	1.1	149,651	2.47	16,093
14	1,080,259	2491	825090	4	20.02	16.77	20.8	17,328,970	17.16	17,324,464
17	836,409	2493	457719	2	43.13;	43.07;	43.1			
					43.28	43.31	3	32,764,035	43.07	33,363,987
18	756,520	2495	602437	3	34.59	34.38	34.9			
							3	18,194,245	32.69	17,326,166
20	578,975	2570	442527	4	26.08	22.83	26.3			
							3	30,625,333	22.83	30,471,508
26	404,561	2571	305538	2	34.13	33.78	34.2			
							3	17,090,936	33.78	17,105,836
31	539,176	2640	522003	5	54.11	60.53	54.3			
							5	46,751,901	60.43	46,184,112
32	413,828	2641	21624	9	61.67	60.85	61.6			
							7	37,693,815	60.85	37,809,922
41	249,078	2688	143205	8	3.80	3.50	3.62	6,606,892	3.5	6,462,704
54	193,117	2758	135336	3	44.71		44.7			
							1	32,560,216		
57	168,045	2759	6015	8	109.07	108.76	108.			
							6	65,161,200	108.8	65,260,200
71	152,258	2905	8348	3	0	0	0	3,178,533	0	3,178,533
164	60,483	2492	448098	2	34.38	34.02	34.7	20,821,800	34.3	20,577,000

Table S4.12 List of genes underlying loci with a maximum segregation distortion ratio. Their location on the BRASUZ1 genome and their annotation and associated reference paper are provided. (likage group: LG, scaffold id: Sca, locus name: AT ID)





Parent	LG	Sca	Location (bp)	AT ID	Function or putative function
<i>E.g</i>	1	1	12,741,837	AT5G54160.1	Atomt1, Caffeate O-Methyltransferase 1, Comt1, O-Methyltransferase 1, Omt1
<i>E.g</i>	1	1	12,854,560	AT2G38560.1	Rdo2, Reduced Dormancy 2, Tfiis, Transcript Elongation Factor Iis
<i>E.g</i>	1	2	50,454,931	AT5G01030.1	Protein Of Unknown Function (DUF3527)
<i>E.g</i>	1	1	17,934,261	AT1G16670.1	Protein Kinase Superfamily Protein
<i>E.g</i>	1	1	26,349,566	AT3G09410.1	Pectinacetylesterase Family Protein
<i>E.g</i>	1	1	24,667,508	AT3G53000.1 *	Atp2-A15, Phloem Protein 2-A15, Pp2-A15
<i>E.g</i>	2	2	18,176,413	AT5G19040.1	Arabidopsis Thaliana Isopentenyltransferase 5, Atipt5, Ipt5, Isopentenyltransferase 5
<i>E.g</i>	2	2	18,334,429	AT3G23600.1	Alpha/Beta-Hydrolases Superfamily Protein
<i>E.g</i>	2	2	18,460,653	AT3G14470.1	NB-ARC Domain-Containing Disease Resistance Protein
<i>E.g</i>	2	2	23,456,565	AT3G03340.1 *	Une6, Unfertilized Embryo Sac 6
<i>E.g</i>	2	2	22,975,758	AT1G60500.1	Drp4c, Dynamin Related Protein 4c
<i>E.g</i>	2	2	37,963,058	AT1G54710.1	Atatg18h, Atg18h, Homolog Of Yeast Autophagy 18 (Atg18) H
<i>E.g</i>	2	2	43,540,743	AT2G30105.1 *	CONTAINS Interpro DOMAIN/S: Leucine-Rich Repeat, Typical Subtype
<i>E.g</i>	2	2	48,118,573	AT1G51110.1	Plastid-Lipid Associated Protein PAP / Fibrillin Family Protein
<i>E.g</i>	2	2	55,373,761	AT1G25550.1	Myb-Like Transcription Factor Family Protein
<i>E.g</i>	3	3	12,918,875	AT1G19250.1	Flavin-Dependent Monooxygenase 1, Fmo1
<i>E.g</i>	3	3	13,656,913	AT5G17050.1	Udp-Glucosyl Transferase 78d2, Ugt78d2
<i>E.g</i>	3	3	15,900,197	AT1G35190.1 *	2-Oxoglutarate (2OG) And Fe(II)-Dependent Oxygenase Superfamily Protein
<i>E.g</i>	3	3	22,044,378	AT5G52580.1	Rabgap/TBC Domain-Containing Protein
<i>E.g</i>	3	3	22,205,082	AT2G10940.1	Bifunctional Inhibitor/Lipid-Transfer Protein/Seed Storage 2S Albumin Superfamily Protein
<i>E.g</i>	3	3	54,190,153	AT3G24350.1	Atsyp32, Syntaxin Of Plants 32, Syp32
<i>E.g</i>	7	7	31,013,349	AT4G39140.1	RING/U-Box Superfamily Protein
<i>E.g</i>	7	7	32,703,677	AT4G38650.1	Glycosyl Hydrolase Family 10 Protein
<i>E.g</i>	7	7	33,149,949	AT4G35140.1	Transducin/WD40 Repeat-Like Superfamily Protein
<i>E.g</i>	7	7	33,053,148	AT4G38510.5 *	Atvab2, V-Atpase B Subunit 2, Vab2
<i>E.g</i>	7	7	33,323,760	AT4G38470.1	Serine/Threonine/Tyrosine Kinase 46, Sty46
<i>E.u</i>	2	2	14,729,054	AT5G43850.1	Ard4, Atard4
<i>E.u</i>	2	2	18,375,373	AT3G23600.1	Alpha/Beta-Hydrolases Superfamily Protein
<i>E.u</i>	2	2	27,161,208	AT4G34640.1	Erg9, Sqs1, Squalene Synthase 1
<i>E.u</i>	2	2	27,263,099	AT4G34640.1	Erg9, Sqs1, Squalene Synthase 1
<i>E.u</i>	2	2	23,831,338	AT1G65540.1 *	Atletm2, Letm2, Leucine Zipper-Ef-Hand-Containing Transmembrane Protein 2
<i>E.u</i>	2	2	28,755,000	AT3G05740.1	RECQ Helicase L1 (RECQI1)
<i>E.u</i>	2	2	40,457,792	AT1G56300.1	Chaperone Dnaj-Domain Superfamily Protein
<i>E.u</i>	2	2	36,270,917	AT3G05510.1	Phospholipid/Glycerol Acyltransferase Family Protein
<i>E.u</i>	2	2	56,597,841	AT1G69230.1	Sp112, Spiral1-Like2
<i>E.u</i>	5	5	2,557,979	AT4G17790.1	SNARE Associated Golgi Protein Family
<i>E.u</i>	5	5	5,101,291	AT1G12740.1	Cytochrome P450, Family 87, Subfamily A,



Parent	LG	Sca	Location (bp)	AT ID	Function or putative function
				*	Polypeptide 2", Cyp87a2
<i>E.u</i>	5	5	42,820,799	AT1G53210.1	Atncl, Na <sup>+</sup> /Ca <sup>2+</sup> Exchanger, Ncl
<i>E.u</i>	5	5	70,194,206	AT1G69630.1	F-Box/RNI-Like Superfamily Protein
<i>E.u</i>	5	5	67,388,620	AT5G47635.1	Pollen Ole E 1 Allergen And Extensin Family Protein
<i>E.u</i>	5	5	73,608,258	AT1G31120.1 *	K <sup>+</sup> Uptake Permease 10, Kup10
<i>E.u</i>	6	6	3,712,164	AT2G26560.1	Patatin-Like Protein 2, Phospholipase A 2a, Pla Iia, Pla2a, Plaii Alpha, Plp2
<i>E.u</i>	6	6	545,682	AT2G26560.1	Patatin-Like Protein 2, Phospholipase A 2a, Pla Iia, Pla2a, Plaii Alpha, Plp2
<i>E.u</i>	6	6	7,502,525	AT2G33030.1	Receptor Like Protein 25 (RLP25)
<i>E.u</i>	6	6	9,706,368	AT1G11910.1 *	Apa1, Aspartic Proteinase A1, Atapa1
<i>E.u</i>	6	6	10,011,705	AT1G50670.1	OTU-Like Cysteine Protease Family Protein
<i>E.u</i>	6	6	19,770,345	AT3G18400.1	Anac058, Nac Domain Containing Protein 58, Nac058
<i>E.u</i>	6	6	20,019,948	AT5G42710.1	Unknown Protein
<i>E.u</i>	6	6	23,913,579	AT1G19900.1	Glyoxal Oxidase-Related Protein
<i>E.u</i>	6	6	24,123,124	AT2G16430.2	Atpap10, Pap10, Purple Acid Phosphatase 10
<i>E.u</i>	6	6	24,703,474	AT5G66460.1	Atman7, Endo-Beta-Mannase 7, Man7
<i>E.u</i>	6	6	24,841,599	AT2G18280.1	Attlp2, Tlp2, Tubby Like Protein 2
<i>E.u</i>	6	6	27,682,407	AT1G55915.1	Zinc Ion Binding
<i>E.u</i>	6	6	28,122,664	AT1G75590.1	SAUR-Like Auxin-Responsive Protein Family
<i>E.u</i>	6	6	28,264,199	AT1G19880.1	Regulator Of Chromosome Condensation (RCC1) Family Protein
<i>E.u</i>	7	7	44,034,491	AT4G35240.1	Protein Of Unknown Function (DUF630 And DUF632)
<i>E.u</i>	10	10	10,076,755	AT2G28350.1	Arf10, Auxin Response Factor 10
<i>E.u</i>	11	11	2,612,993	AT4G25433.1	Peptidoglycan-Binding Lysm Domain-Containing Protein
<i>E.u</i>	11	11	4,769,787	AT5G24430.1	Calcium-Dependent Protein Kinase (CDPK) Family Protein



## References

- Arnold ML, Hamrick JL, Bennett BD. 1993.** Interspecific Pollen Competition and Reproductive Isolation in Iris. *Journal of Heredity* **84**(1): 13-16.
- Benjamini Y, Yekutieli D. 2001.** The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* **29**(4): 1165-1188.
- Bikard D, Patel D, Le Metté C, Giorgi V, Camilleri C, Bennett MJ, Loudet O. 2009.** Divergent Evolution of Duplicate Genes Leads to Genetic Incompatibilities Within *A. thaliana*. *Science* **323**(5914): 623-626.
- Botstein D, White RL, Skolnick M, Davis RW. 1980.** Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics* **32**(3): 314.
- Bradshaw HD, Jr., Stettler RF. 1994.** Molecular genetics of growth and development in Populus. II. Segregation distortion due to genetic load. *Theoretical and Applied Genetics* **89**(5): 551-558.
- Brondani R, Brondani C, Tarchini R, Grattapaglia D. 1998.** Development, characterisation and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *Theoretical and Applied Genetics* **97**: 816 - 827.
- Brondani R, Williams E, Brondani C, Grattapaglia D. 2006.** A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. *BMC Plant Biology* **6**(1): 20.
- Burt DW. 2002.** Comparative mapping in farm animals. *Briefings in Functional Genomics & Proteomics* **1**(2): 159-168.
- Chagné D, Crowhurst RN, Troggio M, Davey MW, Gilmore B, Lawley C, Vanderzande S, Hellens RP, Kumar S, Cestaro A, et al. 2012.** Genome-Wide SNP Detection, Validation, and Development of an 8K SNP Array for Apple. *PLoS ONE* **7**(2): e31745.
- Cheema J, Dicks J. 2009.** Computational approaches and software tools for genetic linkage map estimation in plants. *Briefings in Bioinformatics* **10**(6): 595-608.
- Choi H-K, Mun J-H, Kim D-J, Zhu H, Baek J-M, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, et al. 2004.** Estimating genome conservation between crop and model legume species. *Proceedings of the National Academy of Sciences of the United States of America* **101**(43): 15289-15294.
- Collard B, Mace E, McPhail M, Wenzl P, Cakir M, Fox G, Poulsen D, Jordan D. 2009.** How accurate are the marker orders in crop linkage maps generated from large marker datasets? *Crop and Pasture Science* **60**(4): 362-372.
- Crismani W, Girard C, Mercier R. 2013.** Tinkering with meiosis. *Journal of Experimental Botany* **64**(1): 55-65.
- Dickinson GR, Lee DJ, Wallace HM. 2012.** The influence of pre- and post-zygotic barriers on interspecific *Corymbia* hybridization. *Annals of Botany* **109**(7): 1215-1226.
- Doyle J, Doyle J. 1990.** Isolation of plant DNA from fresh tissue. *Focus* **12**: 13-15.
- Ellis MF, Sedgley M, Gardner JA. 1991.** Interspecific Pollen-Pistil Interaction in *Eucalyptus* L'Hér. (Myrtaceae): The Effect of Taxonomic Distance. *Annals of Botany* **68**(3): 185-194.
- Endelman J. 2011.** New algorithm improves fine structure of the barley consensus SNP map. *BMC Genomics* **12**(1): 407.
- Ferreira A, da Silva MF, Silva L, Cruz CD. 2006.** Estimating the effects of population size and type on the accuracy of genetic maps. *Genetics and Molecular Biology* **29**(1): 187-192.
- Ganal MW, Durstewitz G, Polley A, Bérard A, Buckler ES, Charcosset A, Clarke JD, Graner E-M, Hansen M, Joets J, et al. 2011.** A Large Maize (*Zea mays* L.) SNP Genotyping Array:



- Development and Germplasm Genotyping, and Genetic Mapping to Compare with the B73 Reference Genome. *PLoS ONE* **6**(12): e28334.
- Geraldes A, DiFazio SP, Slavov GT, Ranjan P, Muchero W, Hannemann J, Gunter LE, Wymore AM, Grassa CJ, Farzaneh N, et al. 2013.** A 34K SNP genotyping array for *Populus trichocarpa*: Design, application to the study of natural populations and transferability to other *Populus* species. *Molecular Ecology Resources* **13**(2): 306-323.
- Gion J-M, Carouche A, Deweer S, Bedon F, Pichavant F, Charpentier J-P, Bailleres H, Rozenberg P, Carocha V, Ognouabi N, et al. 2011.** Comprehensive genetic dissection of wood properties in a widely-grown tropical tree: *Eucalyptus*. *BMC Genomics* **12**(1): 301.
- Gion J-M, Rech P, Grima-Pettenati J, Verhaegen D, Plomion C. 2000.** Mapping candidate genes in *Eucalyptus* with emphasis on lignification genes. *Molecular Breeding* **6**(5): 441-449.
- Gion J, Chaumeil P, Plomion C. in prep.** EucaMaps : Linking genetic and QTL maps to the *Eucalyptus grandis* genome.
- Gore PL, Potts BM, Volker PW, Megalos J. 1990.** Unilateral Cross-Incompatibility in *Eucalyptus*: the Case of Hybridisation Between *E. globulus* and *E. nitens*. *Australian Journal of Botany* **38**(4): 383-394.
- Grattapaglia D, Bradshaw Jr. HD. 1994.** Nuclear DNA content of commercially important *Eucalyptus* species and hybrids. *Canadian Journal of Forest Research* **24**(5): 1074-1078.
- Grattapaglia D, Sederoff R. 1994.** Genetic Linkage Maps of *Eucalyptus grandis* and *Eucalyptus urophylla* Using a Pseudo-Testcross Mapping Strategy and RAPD Markers. *Genetics* **137**(4): 1121-1137.
- Grattapaglia D, Silva-Junior O, Kirst M, de Lima B, Faria D, Pappas G. 2011.** High-throughput SNP genotyping in the highly heterozygous genome of *Eucalyptus*: assay success, polymorphism and transferability across species. *BMC Plant Biology* **11**(1): 65.
- Hackett CA, Broadfoot LB. 2003.** Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. *Heredity* **90**(1): 33-38.
- Haldane JBS. 1919.** The combination of linkage values, and the calculation of distance between the loci of linked factors. *J Genet* **8**: 299-309.
- Howe GT, Yu JB, Knaus B, Cronn R, Kolpak S, Dolan P, Lorenz WW, Dean JFD. 2013.** A SNP resource for Douglas-fir: de novo transcriptome assembly and SNP detection and validation. *BMC Genomics* **14**.
- Hudson CJ, Freeman JS, Kullán ARK, Petroli CD, Sansaloni CP, Kilian A, Detering F, Grattapaglia D, Potts BM, Myburg AA, et al. 2012a.** A reference linkage map for eucalyptus. *BMC Genomics* **13**(1): 240.
- Hudson CJ, Kullán ARK, Freeman JS, Faria DA, Grattapaglia D, Kilian A, Myburg AA, Potts BM, Vaillancourt RE. 2012b.** High synteny and colinearity among *Eucalyptus* genomes revealed by high-density comparative genetic mapping. *Tree Genetics & Genomes* **8**(2): 339-352.
- Jackson BN, Schnable PS, Aluru S. 2008.** Consensus Genetic Maps as Median Orders from Inconsistent Sources. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **5**(2): 161-171.
- Kaul S, Koo HL, Jenkins J, Rizzo M, Rooney T, Tallon LJ, Feldblyum T, Nierman W, Benito MI, Lin XY, et al. 2000.** Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**(6814): 796-815.
- Kent WJ. 2002.** BLAT—The BLAST-Like Alignment Tool. *Genome Research* **12**(4): 656-664.
- Kole C. 2007.** *Genome Mapping and Molecular Breeding in Plants: Forest trees*. Berlin: Springer.
- Koornneef M, Vaneden J, Hanhart CJ, Stam P, Braaksma FJ, Feenstra WJ. 1983.** Linkage map of *arabidopsis thaliana*. *Journal of Heredity* **74**(4): 265-272.



- Krutovsky KV, Troggio M, Brown GR, Jermstad KD, Neale DB. 2004.** Comparative Mapping in the Pinaceae. *Genetics* **168**(1): 447-461.
- Kullan ARK, van Dyk M, Jones N, Kanzler A, Bayley A, Myburg A. 2012.** High-density genetic linkage maps with over 2,400 sequence-anchored DArT markers for genetic dissection in an F2 pseudo-backcross of *Eucalyptus grandis* × *E. urophylla*. *Tree Genetics & Genomes* **8**(1): 163-175.
- Kumar S, Banks TW, Cloutier S. 2012.** SNP Discovery through Next-Generation Sequencing and Its Applications. *International Journal of Plant Genomics* **2012**: 15.
- Lander ES, Int Human Genome Sequencing C, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, et al. 2001.** Initial sequencing and analysis of the human genome. *Nature* **409**(6822): 860-921.
- Li H, Durbin R. 2009.** Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**(14): 1754-1760.
- Mackay TFC, Stone EA, Ayroles JF. 2009.** The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet* **10**(8): 565-577.
- Margarido GRA, Souza AP, Garcia AAF. 2007.** OneMap: software for genetic mapping in outcrossing species. *Hereditas* **144**(3): 78-79.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010.** The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**(9): 1297-1303.
- Mollinari M, Margarido GRA, Vencovsky R, Garcia AAF. 2009.** Evaluation of algorithms used to order markers on genetic maps. *Heredity* **103**(6): 494-502.
- Myburg AA, Griffin AR, Sederoff RR, Whetten RW. 2003.** Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. *Theoretical and Applied Genetics* **107**(6): 1028-1042.
- Myburg AA, Vogl C, Griffin AR, Sederoff RR, Whetten RW. 2004.** Genetics of Postzygotic Isolation in *Eucalyptus*: Whole-Genome Analysis of Barriers to Introgression in a Wide Interspecific Cross of *Eucalyptus grandis* and *E. globulus*. *Genetics* **166**(3): 1405-1418.
- Neale DB, Kremer A. 2011.** Forest tree genomics: growing resources and applications. *Nat Rev Genet* **12**(2): 111-122.
- Nelson JC 2005.** Methods and Software for Genetic Mapping. *The Handbook of Plant Genome Mapping*: Wiley-VCH Verlag GmbH & Co. KGaA, 53-74.
- Neves L, MC Mamani E, Alfenas A, Kirst M, Grattapaglia D. 2011.** A high-density transcript linkage map with 1,845 expressed genes positioned by microarray-based Single Feature Polymorphisms (SFP) in *Eucalyptus*. *BMC Genomics* **12**(1): 189.
- Novaes E, Drost DR, Farmerie WG, Pappas GJ, Grattapaglia D, Sederoff RR, Kirst M. 2008.** High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics* **9**.
- Os H, Stam P, Visser RF, Eck H. 2005.** RECORD: a novel method for ordering loci on a genetic linkage map. *Theoretical and Applied Genetics* **112**(1): 30-40.
- Pavy N, Gagnon F, Rigault P, Blais S, Deschenes A, Boyle B, Pelgas B, Deslauriers M, Clement S, Lavigne P, et al. 2013.** Development of high-density SNP genotyping arrays for white spruce (*Picea glauca*) and transferability to subtropical and nordic congeners. *Molecular Ecology Resources* **13**(2): 324-336.
- Peirce JL, Broman KW, Lu L, Williams RW. 2007.** A Simple Method for Combining Genetic Mapping Data from Multiple Crosses and Experimental Designs. *PLoS ONE* **2**(10): e1036.





- Petroli CD, Sansaloni CP, Carling J, Steane DA, Vaillancourt RE, Myburg AA, da Silva OB, Jr., Pappas GJ, Jr., Kilian A, Grattapaglia D. 2012. Genomic Characterization of DArT Markers Based on High-Density Linkage Analysis and Physical Mapping to the *Eucalyptus* Genome. *PLoS ONE* 7(9): e44684.
- Potts BM, Dungey HS. 2004. Interspecific hybridization of *Eucalyptus*: key issues for breeders and geneticists. *New Forests* 27(2): 115-138.
- Praça MM, Carvalho CR, Novaes CRDB. 2009. Nuclear DNA content of three *Eucalyptus* species estimated by flow and image cytometry. *Australian Journal of Botany* 57(6): 524-531.
- Price AH. 2006. Believe it or not, QTLs are accurate! *TRENDS in Plant Science* 11(5): 213-216.
- Rahmé J, Widmer A, Karrenberg S. 2009. Pollen competition as an asymmetric reproductive barrier between two closely related *Silene* species. *Journal of Evolutionary Biology* 22(9): 1937-1943.
- Rieseberg LH, Carney SE. 1998. Plant hybridization. *New Phytologist* 140(4): 599-624.
- Ritter E, Gebhardt C, Salamini F. 1990. Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. *Genetics* 125(3): 645-654.
- Rocha RB, Barros EG, Cruz CD, Rosado AM, Araujo EFd. 2007. Mapping of QTLs related with wood quality and developmental characteristics in hybrids ( *Eucalyptus grandis* \* *Eucalyptus urophylla*). *Revista Arvore* 31(1): 13-24.
- Ronin Y, Mester D, Minkov D, Belotserkovski R, Jackson BN, Schnable PS, Aluru S, Korol A. 2012. Two-Phase Analysis in Consensus Genetic Mapping. *G3: Genes/Genomes/Genetics* 2(5): 537-549.
- Ronin Y, Mester D, Minkov D, Korol A. 2010. Building reliable genetic maps: different mapping strategies may result in different maps. *Nat. Science* 2: 576-589.
- Salvi S, Tuberosa R. 2007. Cloning QTLs in Plants. In: Varshney R, Tuberosa R eds. *Genomics-Assisted Crop Improvement*: Springer Netherlands, 207-225.
- Sansaloni CP, Petroli CD, Carling J, Hudson CJ, Steane DA, Myburg AA, Grattapaglia D, Vaillancourt RE, Kilian A. 2010. A high-density Diversity Arrays Technology (DArT) microarray for genome-wide genotyping in *Eucalyptus*. *Plant Methods* 6.
- Schlotterer C. 2004. The evolution of molecular markers - just a matter of fashion? *Nat Rev Genet* 5(1): 63-69.
- Shen R, Fan J-B, Campbell D, Chang W, Chen J, Doucet D, Yeakley J, Bibikova M, Wickham Garcia E, McBride C, et al. 2005. High-throughput SNP genotyping on universal bead arrays. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 573(1-2): 70-82.
- Slavov GT, DiFazio SP, Martin J, Schackwitz W, Muchero W, Rodgers-Melnick E, Lipphardt MF, Pennacchio CP, Hellsten U, Pennacchio LA, et al. 2012. Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytologist* 196(3): 713-725.
- Stam P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: Join Map. *The Plant Journal* 3(5): 739-744.
- Steane DA, Nicolle D, McKinnon GE, Vaillancourt RE, Potts BM. 2002. Higher-level relationships among the eucalypts are resolved by ITS-sequence data. *Australian Systematic Botany* 15(1): 49-62.
- Steane DA, Nicolle D, Sansaloni CP, Petroli CD, Carling J, Kilian A, Myburg AA, Grattapaglia D, Vaillancourt RE. 2011. Population genetic analysis and phylogeny reconstruction in *Eucalyptus* (Myrtaceae) using high-throughput, genome-wide genotyping. *Molecular Phylogenetics and Evolution* 59(1): 206-224.



- Sturtevant AH. 1913.** The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *Journal of Experimental Zoology* **14**(1): 43-59.
- Thamarus KA, Groom K, Murrell J, Byrne M, Moran GF. 2002.** A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre, and floral traits. *Theoretical and Applied Genetics* **104**(2-3): 379-387.
- Törjék O, Witucka-Wall H, Meyer R, Korff M, Kusterer B, Rautengarten C, Altmann T. 2006.** Segregation distortion in *Arabidopsis* C24/Col-0 and Col-0/C24 recombinant inbred line populations is due to reduced fertility caused by epistatic interaction of two loci. *Theoretical and Applied Genetics* **113**(8): 1551-1561.
- Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, et al. 2006.** The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**(5793): 1596-1604.
- Van Ooijen JW 2011.** JoinMap 4.1, Software for the Calculation of Genetic Linkage Maps in Experimental Populations. Wageningen, Netherlands: Kyazma BV.
- Wu Y, Close TJ, Lonardi S. 2011.** Accurate construction of consensus genetic maps via integer linear programming. *IEEE/ACM Trans Comput Biol Bioinform* **8**(2): 381-394.
- Wurschum T. 2012.** Mapping QTL for agronomic traits in breeding populations. *Theoretical and Applied Genetics* **125**(2): 201-210.
- Yelina NE, Choi K, Chelysheva L, Macaulay M, de Snoo B, Wijnker E, Miller N, Drouaud J, Grelon M, Copenhaver GP, et al. 2012.** Epigenetic Remodeling of Meiotic Crossover Frequency in *Arabidopsis thaliana* DNA Methyltransferase Mutants. *PLoS Genet* **8**(8): e1002844.
- Yin T, DiFazio SP, Gunter LE, Zhang X, Sewell MM, Woolbright SA, Allan GJ, Kelleher CT, Douglas CJ, Wang M, et al. 2008.** Genome structure and emerging evidence of an incipient sex chromosome in *Populus*. *Genome Research* **18**(3): 422-430.
- Yin TM, DiFazio SP, Gunter LE, Riemenschneider D, Tuskan GA. 2004.** Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map. *Theoretical and Applied Genetics* **109**(3): 451-463.
- Zhang X-L, Li X-X, Ni L-Y, Guo Y-H. 2011.** When interspecific prezygotic barriers break down: hybridization between two *Potamogeton* species (*P. wrightii* and *P. perfoliatus*) and comparison of the artificial and natural hybrids. *Plant Systematics and Evolution* **295**(1-4): 119-128.





## **5 Genetic architecture of carbon isotope composition and growth in *Eucalyptus***



## **Genetic architecture of carbon isotope composition and growth in *Eucalyptus***

Jérôme BARTHOLOME<sup>1, 2, 3</sup>, André MABIALA<sup>4</sup>, Bruno SAVELLI<sup>5</sup>, Didier BERT<sup>2, 3</sup>, Oliver BRENDEL<sup>6, 7</sup>, Christophe PLOMION<sup>2, 3</sup>, Jean-Marc GION<sup>1, 2, 3\*</sup>

<sup>1</sup> CIRAD, UMR AGAP, F-33612 Cestas, France

<sup>2</sup> INRA, UMR BIOGECO, F-33612 Cestas, France

<sup>3</sup> Univ. Bordeaux, UMR 1202, F-33610 Cestas, France

<sup>4</sup> CRDPI, BP 1291, Pointe Noire, Rep. of Congo

<sup>5</sup> CNRS, UMR 5546, BP 42617, F-31326 Castanet-Tolosan, France

<sup>6</sup> INRA, UMR 1137, F-54280 Champenoux, France

<sup>7</sup> Nancy Université, UMR 1137, F-54506 Vandoeuvre-les-Nancy, France

\* Corresponding author: jean-marc.gion@cirad.fr



## 5.1 Introduction

Planted forests face two main interrelated issues: the need to respond to an increasing demand for wood biomass (FAO, 2006) and to maintain production under increasing environmental constraints (IPCC, 2007). Among the consequences of climate change, water limitation caused by changes in rainfall regimes coupled with the increase in average temperatures, will likely be a major limiting factor for the productivity of planted forests by reducing tree growth (Pereira & Pallardy, 1989) or by causing the death of trees in the case of severe drought (Breda *et al.*, 2006; Allen *et al.*, 2010; Phillips *et al.*, 2010). For these reasons, water consumption by forest trees has been widely studied at tree (Wullschlegel *et al.*, 1998), stand (Whitehead & Beadle, 2004; Almeida *et al.*, 2007; Stape *et al.*, 2008; Hubbard *et al.*, 2010), and catchment scales (Scott & Prinsloo, 2008; Silveira & Alonso, 2009). This knowledge is a prerequisite for optimising silvicultural strategies (Battaglia *et al.*, 2009; White *et al.*, 2009) as well as for genetic improvement of trees to cope with water shortage (Pita *et al.*, 2005).

At the plant scale, the ratio of the amount of harvested biomass to the amount of water transpired to produce it is termed water use efficiency (WUE) (Dye, 2000; Condon *et al.*, 2004). This trait is directly related to light absorption, which drives canopy photosynthesis and transpiration, to leaf photosynthetic capacities and to stomatal conductance which is driven by tree and leaf hydraulic characteristics (Wullschlegel *et al.*, 1998; Whitehead & Beadle, 2004). At an instantaneous scale, intrinsic WUE can be defined as the ratio of  $\text{CO}_2$  assimilation ( $A$ ) to stomatal conductance for water ( $g_s$ ). Intrinsic WUE was found to be linked to the composition of stable carbon isotopes ( $\delta^{13}\text{C}$ ) of plant biomass by Farquhar *et al.* (1984; 1989). Indeed,  $\delta^{13}\text{C}$  has been used as an integrative and robust indicator of WUE. Thanks to high throughput characterization of  $\delta^{13}\text{C}$  with mass spectrometry, the genetic variability and architecture of carbon isotope composition has been widely studied within species in crop plants (e.g. *Glycine max* (Specht *et al.*, 2001), *Hordeum vulgare* (Teulat *et al.*, 2002) *Oryza sativa* (Ishimaru *et al.*, 2001) *Triticum aestivum* (Rebetzke *et al.*, 2008)) and long-lived species (e.g. *Castanea sativa* (Casasoli *et al.*, 2004), *Pinus pinaster* (Brendel *et al.*, 2002), *Populus* spp. (Monclus *et al.*, 2012), *Quercus robur* (Brendel *et al.*, 2008) and *Vitis vinifera* x *riparia* (Marguerit *et al.*, 2012)). In relation with a small genotype-by-environment interaction and moderate to high heritability (Hubick *et al.*, 1988; Condon & Richards, 1992; Rebetzke *et*



*al.*, 2002; Lambrides *et al.*, 2004; Comstock *et al.*, 2005), some stable genomic regions across environments were detected including in *Oryza sativa* (Laza *et al.*, 2006), *Triticum aestivum* (Rebetzke *et al.*, 2008), *Quercus robur* (Brendel *et al.*, 2008) and *Pinus pinaster* (Marguerit *et al.*, 2014). Together with increased genome sequencing, these studies open new opportunities for the molecular characterisation of carbon isotope composition and indirectly WUE.

Forward genetics approaches have been developed to investigate the genes underlying QTLs for carbon isotope composition (Masle *et al.*, 2005). In woody species, colocalization between candidate genes and QTLs in mapping pedigrees (Marguerit *et al.* 2012) or the association between polymorphism in candidate genes and trait variation in populations of unrelated trees (Gonzalez-Martinez *et al.*, 2008; Cumbie *et al.*, 2011) have been used to identify putative genes or polymorphisms involved in trait variation. However, this strategy has been limited by either lack of mapping accuracy owing to the use of low density genetic maps (in linkage mapping approaches) or the drop in linkage disequilibrium in a few hundred bp (in association mapping studies). Thanks to the availability of genome sequences and technological advances which enable the production of high density gene-based maps and their alignment with reference genomes, it has now become possible to mine more precisely for positional candidate genes underlying trait-QTLs without prior information. This approach, based on the analysis of GO term enrichment, has been used in grapevine for berry related traits (Costantini *et al.*, 2008) and in poplar for several quantitative traits related to growth (Monclus *et al.*, 2012). The main aim of enrichment analysis is to assign biological meaning to a group of genes using their gene ontology (GO). The aim of GO is to standardise the representation of genes and gene product attributes across species and databases by assigning a set of associations from controlled vocabulary of terms to specific genes, either chosen by trained curators (<http://www.geneontology.org/>) or generated automatically. In the context of polygenic traits such as those used in QTL approaches, enrichment analysis of GO terms appears to be an interesting approach to decipher the metabolic processes underlying (i) one QTL possibly involved in a cluster of closely linked genes, or (ii) several QTLs over the genome for a particular trait.

In the case of *Eucalyptus*, the most widely planted hardwood in the world (FAO, 2007), breeding programmes have focused on a limited number of fast growing *Eucalyptus* species and hybrids (Eldridge *et al.*, 1993; Grattapaglia, 2007; Grattapaglia & Kirst, 2008) leading to



highly productive plantations (Stape *et al.*, 2010; Gonçalves *et al.*, 2013). As high growing rates are often related to high rates of water use, water consumption by *Eucalyptus* has been widely study at the stand (Whitehead & Beadle, 2004; Stape *et al.*, 2010). However, there is no report of the genetic architecture of water use efficiency and its links with growth in *Eucalyptus*. The recent publication of the *Eucalyptus* BASUZ1 reference genome, has made it possible to investigate the genes underlying  $\delta^{13}\text{C}$  QTLs.

In this context, the aims of the present study were to (i) analyse the phenotypic variability of  $\delta^{13}\text{C}$  at different levels of integration (seasonal and rotation) and its relation with growth, (ii) characterise the genetic architecture of  $\delta^{13}\text{C}$  and its interplay with the environment and tree growth and (iii) propose putative candidate genes underlying the QTLs detected for  $\delta^{13}\text{C}$  using two high resolution genetic maps anchored on the *Eucalyptus* BRASUZ1 reference genome combined with a "digital candidate gene approach" based on GO term enrichment analysis.

## 5.2 Materials and Methods

### 5.2.1 Mapping population and field experiments

The mapping population was generated from an interspecific cross between *Eucalyptus urophylla* (*E.u.*, female) and *E. grandis* (*E.g.*, male). These two *Eucalyptus* species are known to have contrasted performance in Congolese conditions (Vigneron, 1992). This F1 progeny was established in three different trials using distinct genotypes in each trial. The first trial was planted in April 1993 (P93) and included 201 genotypes, the second in April 1997 (P97) with 190 genotypes and the third in November 2010 (P10) with 960 genotypes. Trials P93 and P97 have been described previously (Verhaegen *et al.*, 1997; Gion *et al.*, 2011; Bartholomé *et al.*, 2013). The density was 667 trees/ha for P93 and P97, and 802 trees/ha for P10. All the trials were planted using a single tree plot design with a triple border row.

### 5.2.2 Environmental characterisation

The three trials were planted at the Kissoko forestry station (4.45° S, 12° E) in the coastal region of Pointe-Noire (Republic of Congo). The soil in this region (Ferralic Arenosols, FAO



$\delta^{13}\text{C}$				↓			↓	
Height		↓		↓			↓	↓
Diameter		↓	↓	↓		↓	↓	↓
Age in months	1	3	6	10	15	18	22	23

Figure 5.1: Timing of leaf sampling for  $\delta^{13}\text{C}$  and growth measurements in the P10 trial. The dry season (DS) is in orange and the rainy season (RS) in blue.



classification) is mainly composed of sand ( $\approx 90\%$ ) with low storage capacity for either water or nutrients (Mareschal *et al.*, 2011). The climate is tropical savanna (Kottek *et al.*, 2006), with two contrasted seasons in terms of water availability: a dry season (DS) from June to September with almost no rain and a rainy season (RS) from October to April with about 1,200 mm of rainfall. Atmospheric humidity and temperatures are generally high (85% and  $25^\circ\text{C}$ ) with low inter-seasonal variations: 2% and  $3^\circ\text{C}$ , respectively.

In the P93 and P97 trials, temperature and rainfall were measured at Pointe-Noire, 25 km from the field trials, which allowed us to estimate the monthly de Martonne aridity index ( $I_{\text{DM}}$ ). As already reported by Bartholomé *et al.* (2013), the  $I_{\text{DM}}$  revealed four additional months of drought ( $I_{\text{DM}} < 15$ ) in the P93 trial compared to the P97 trial in the five first years of growth.

In the P10 trial, climatic data, temperature and relative humidity (CS215, Campbell Scientific), rainfall (ARG100, Campbell Scientific), global solar radiation (CS300, Campbell Scientific) and soil water content (107, Campbell Scientific) at four depths (0.15, 0.5, 1 and 2 m) were recorded directly at the trial site. A period of 30 days before leaf sampling for  $\delta^{13}\text{C}$  measurements was used to characterise the climate experienced by leaves during their formation (Osório *et al.*, 1998).

### 5.2.3 Phenotypic measurements

#### 5.2.3.1 Carbon isotope discrimination

Carbon isotopic discrimination was measured in different compartments of the tree depending on the trial. In the P93 and P97 trials, the measurements were performed on the holocellulose extracted from 1/8 of a wood disc extracted from the full radius to avoid any year effects in the isotopic measurement. These discs were collected at mid-height at 59 months (March 1998) in the P93 trial (115 trees sampled) and at 1.5 m above the ground at 93 months (January 2005) in the P97 trial (142 trees sampled). In the P10 trial, five mature leaves that had just reached full expansion were sampled at the top of canopy. Sampling was done at two different periods: at the end of the dry season (DS) in September 2011, and at the end of the rainy season (RS) in May 2012, on 212 and 219 trees, respectively (Figure 5.1).





Technically, the wood samples (P93 and P97) were ground in a mill and purified to holocellulose using the Jayme-Wise method (Leavitt & Danzer, 1993). The wood powders were placed in a Teflon poach and sequentially treated in a Soxhlet extractor with 2:1 toluene:ethanol, then 100% ethanol to remove lipids. The samples were then boiled in water to remove soluble carbohydrates, bleached at a temperature of 70 °C in an acetic acid acidified solution to which sodium chlorite was added to decompose the lignin. The extraction yielded an average of 79% of holocellulose ( $n = 257$ ,  $\sigma = 0.04\%$ ), which is in agreement with the lignin and lipophilic extractive content of different eucalyptus species: 0.3-0.6% of lipophilic extractives and 18.7-22.6% of Klason lignin (Rencoret *et al.*, 2007). The leaf samples (P10) were dried at 60 °C for 48 h and crushed to a fine power. For all samples, 1 mg of holocellulose or leaf powder was placed in a tin capsule and isotopic composition was measured with a continuous flow elemental analyser (Carlo Erba NA 1500; CE Instruments, Rodano, Italy) coupled with an isotope ratio mass spectrometer (Thermo-Finnigan; Delta S, Bremen, Germany). The carbon isotope composition was calculated relative to the international standard (Pee Dee Belemnite, PDB) using the following equation:

$$\delta^{13}\text{C} = \left( \frac{(^{13}\text{C}/^{12}\text{C})_{\text{sample}}}{(^{13}\text{C}/^{12}\text{C})_{\text{PDB}}} - 1 \right) \times 1000$$

The precision of spectrometric analysis (standard deviation in  $\delta^{13}\text{C}$ ) was assessed with internal laboratory reference material composed of holocellulose of *Pinus pinaste* ( $n=12$ ,  $\delta^{13}\text{C}$  standard deviation=0.149‰).

### 5.2.3.2 Tree growth

In the P93 and P97 trials, height (HT) and circumference (C) are described in Bartholomé *et al.* (2013). Briefly, P93 trees were measured five times up to the age of 59 months and P97 trees were measured 17 times up to the age of 62 months. In the P10 trial, total height (HT) was measured when the trees were 3, 10, 18 and 22 months old and diameter (D) at 0.2 m above soil level was measured at the same ages with additional measurements at 6 and 15 months old (Figure 5.1). In the P10 trial, growth was measured at the end of the DS and RS. Growth increments between two consecutive measurements were calculated as follows:  $\text{Trait } n_{1n_2} = \text{Trait } n_2 - \text{Trait } n_1$ , where Trait  $n$  is the cumulative HT or D at  $n$  months old.



### 5.2.4 Linkage map construction

For each trial, one genetic map per parental genotype was obtained using the two-way pseudo-test cross mapping strategy (Grattapaglia & Sederoff, 1994). For the P93 and P97 trials, linkage analysis was performed using JoinMap® version 4.1 (Stam, 1993; Van Ooijen, 2011). Common markers between P93 and P97 were used to build a consensus genetic map for each parental genotype with Mergemap (Wu *et al.*, 2008; Wu *et al.*, 2011). For P10, the R package *onemap* v2.0-3 (Margarido *et al.*, 2007; Mollinari *et al.*, 2009) was used to construct high resolution genetic maps for *E.g* and *E.u* thanks to the genotyping of 1,025 individuals with a 6K SNP array (Bartholomé *et al.*, 2014). Among the 1,025 individuals, 60 individuals from P93 were also genotyped using the same SNP array, making it possible to link the consensus maps (P93 and P97) and P10 maps. The characteristics of each genetic map are listed in Table S5.1.

### 5.2.5 QTL analysis

The R package *qtl* v1.28-19 (Broman *et al.*, 2003) was used to perform composite interval mapping (Jansen & Stam, 1994; Zeng, 1994) using the Haley Knott regression (Haley & Knott, 1992). A step interval of 1 cM with a genotyping error rate of 0.01 was applied for the consensus maps (P93 and P97). These values were reduced to 0.5 cM and 0.001 for the high resolution maps (P10). For all traits, 10,000 permutations were performed to determine the genome-wide type I error rate of 5% (Churchill & Doerge, 1994). The confidence interval for each significant QTL was calculated using Bayes credible intervals (BCI), at a confidence level of 0.95 (Manichaikul *et al.*, 2006). MapChart (Voorrips, 2002) was used to represent genetic maps and the location of QTLs.

For each  $\delta^{13}\text{C}$ -QTL detected by composite interval mapping, QTL-by-environment interaction (QTL x E) was tested between trials P93 and P97 using markers common to both trials and between seasons (RS and DS for P10). A two-way analysis of variance was performed using the following model:

$$y = \mu + E + M + E \times M + \varepsilon$$

where  $y$  is the phenotypic measurement of each tree,  $\mu$  is the general mean,  $E$  is the environmental effect (P93 vs. P97 or DS vs. RS),  $M$  is the genotypic effect (closest marker



Table 5.1: Climatic variables recorded in the P10 trial over the dry (DS) and rainy (RS) seasons. Data correspond to the average computed over a 30-day period before leaf sampling for  $\delta^{13}\text{C}$  measurement, as well as over the whole season.

	30 days		Whole season	
	DS	RS	DS	RS
TC	21.9	25	21.9	24.9
RH	90.9	95.1	90.7	93.2
VPD	1028	935	1,026	1129
RG	14.3	15.3	14.1	18
PP	3.8	131.6	20.4	846
SWC_15	0.046	0.093	0.047	0.077
SWC_50	0.05	0.084	0.053	0.074
SWC_100	0.06	0.093	0.071	0.084
SWC_200	0.075	0.085	0.082	0.086

Temperature in Celsius (TC), relative humidity in per cent (RH), vapour pressure deficit in Pa (VPD), rainfall in mm (PP), solar irradiance in MJ/m<sup>2</sup>/d (RG) and soil water content in m<sup>3</sup>/m<sup>3</sup> (SWC) measured at depths of 50, 100 and 200 cm.



linked to the QTL),  $E \times M$  is the interaction between the two main effects. Both  $E$  and  $M$  are treated as fixed effects,  $\varepsilon$  is the residual of the model  $\sim N(0, \sigma_\varepsilon^2)$ .

### 5.2.6 Functional enrichment analysis in $\delta^{13}\text{C}$ QTL-regions

It has been shown in different studies that genes controlling a QTL are often closely related to the maximum LOD score (Price, 2006). Therefore, we defined a  $\delta^{13}\text{C}$  QTL-region for gene enrichment analysis as the genetic segment defined by the two adjacent markers of the closest marker to the LOD peak. This analysis was only performed on the high resolution linkage maps obtained from the analysis of the P10 dataset. These QTL-regions were projected onto the *Eucalyptus* genome using the physical location of the mapped markers. For each QTL-region, gene ontology (GO) term enrichment analysis was performed against the whole genome using Fisher's exact test ( $P_{\text{adjusted}} < 0.05$ , (Benjamini & Hochberg, 1995). The different steps of GO analyses were performed using Blast2GO (Conesa *et al.*, 2005). To this end, the 46,315 predicted protein sequences of the *Eucalyptus* genome (version 1.1) were obtained from Phytozome v9.1, (<http://www.phytozome.net/eucalyptus.php>). Blast2GO software was used with default parameters (blastp, Blastdb: nr; eValue: 1E-10, Best 20) to annotate each protein (Annex, Interproscan, Enzymes, Kegg and Go-Slims steps). A total of 24,998 gene IDs of the *Eucalyptus* genome (distributed on the 11 main scaffolds) were successfully annotated with GO terms covering the three main domains: biological process (P), molecular function (F), and cellular component (C).

## 5.3 Results

### 5.3.1 Climatic conditions

In the P10 trial, the two 30-day periods before leaf sampling (in the dry season DS and the rainy season RS) differed in terms of climate (Table 5.1). The DS had a lower average temperature (21.9 °C) and relative humidity (90.9%) than the RS (25 °C and 95.1%); consequently, the DS had a higher vapour pressure deficit (1028 Pa) than the RS (935 Pa). Almost no rainfall was recorded during the DS (3.8 mm) contrary to in the RS (131.6 mm). The solar irradiance was slightly higher in the RS (15.3 MJ/m<sup>2</sup>/d) than in the DS (14.3). Soil water content (SWC) in the DS at all depths (0.15 m: 0.046, 0.5 m: 0.05, 1 m: 0.06 and 2 m 0.075 m<sup>3</sup>/m<sup>3</sup>) was lower than in the RS (0.15 m: 0.093, 0.5 m: 0.084, 1 m: 0.09 and 2 m 0.085

Table 5.2: Summary of  $\delta^{13}\text{C}$  data for the three trials: P93, P97 and P10.

Trial	Season	N	Min	Max	Mean	SD
P93	-	115	-29	-26.3	-27.7	0.24
P97	-	142	-29.5	-26.4	-27.7	0.29
P10	DS	212	-31.5	-27.5	-29.9	0.71
P10	RS	219	-30.2	-27.5	-29	0.47

N: number of trees, Min: minimum, Max: maximum, mean, SD: standard deviation. DS: dry season, RS: rainy season.

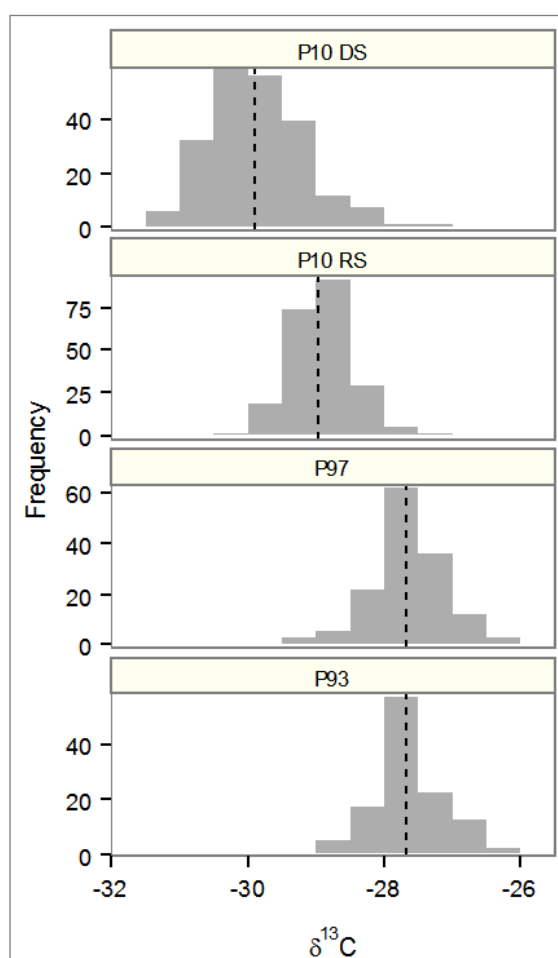


Figure 5.2: Distribution of  $\delta^{13}\text{C}$  (‰) measurements and mean (vertical dashed line) in the three trials (P93, P97 and P10). DS: dry season, RS: rainy season.



$\text{m}^3/\text{m}^3$ ). The climate in the DS was therefore more restricted in terms of soil water availability and atmospheric water demand. These differences in climatic conditions during the two periods of 30 days before leaf sampling were representative of those observed over the entire two seasons (Table 5.1).

### 5.3.2 Trait variation and correlations

$\delta^{13}\text{C}$  and growth traits in the three trials approximated normal distribution (Kolmogorov-Smirnov  $p > 0.05$ ). The range of variation in  $\delta^{13}\text{C}$  was similar in P93 (2.7‰) and P97 (3.1‰), whereas in P10, the range of variation was higher in the DS (4‰) than in the RS (2.7‰) (Table 5.2).  $\delta^{13}\text{C}$  was higher on average in the P10 samples (-29.9‰ and -29‰) than in the P93 and P97 samples (-27.7‰, Table 5.2, Figure 5.2). This difference could be related to the compartments studied (leaves in the P10 trial vs. wood in the P93 and P97 trials) but also to the time scale of measurements (seasonal for P10 vs. rotation duration for P93 and P97). In the P10 trial, a significant difference in  $\delta^{13}\text{C}$  (0.9‰) was found between the DS (-29.9‰) and the RS (-29‰) showing a strong season effect. But surprisingly,  $\delta^{13}\text{C}$  in the DS was lower than in the RS. Additionally, the correlation for  $\delta^{13}\text{C}$  between the two seasons was significant but weak ( $\rho = 0.19$ ,  $p < 0.05$ , Table S5.2) highlighting a change in genotype ranking for  $\delta^{13}\text{C}$  between the two study periods (DS and RS).

In the P10 trial, growth traits presented marked variations (CV<sub>p</sub> [0.13; 0.41], Table S5.2). In all three trials,  $\delta^{13}\text{C}$  presented almost non-significant or weak correlations with growth traits ( $-0.3 < \rho < 0.3$ ; Table S5.3). In the P93 and P97 trials, no significant trend was found in the correlations between  $\delta^{13}\text{C}$  and growth traits. Considering  $\delta^{13}\text{C}$  in the leaves (P10), no correlation was found with growth traits in the RS. Conversely, 13 correlations out of 20 pairwise comparisons were significant in the DS even if the coefficients of correlation did not exceed 0.3. These results suggest a loose relationship between tree size and  $\delta^{13}\text{C}$  during the DS which did not hold at a larger time scale in the P93 and P97 trials.

### 5.3.3 QTL for $\delta^{13}\text{C}$ and growth traits

#### 5.3.3.1 Genetic architecture of $\delta^{13}\text{C}$

QTLs for  $\delta^{13}\text{C}$  were found for all trials and seasons (Table 5.3): five QTLs for P93, two for P97 and five for P10 (three for DS and four for RS). These QTLs were detected on the two


Table 5.3: QTLs for  $\delta^{13}\text{C}$  detected by composite interval mapping in all three trials (P93, P97 and P10).

Parent	Trial	Season	n	LG	Position	LOD	BCI	PEV	PEV tot	ASE
<i>E.g</i>	P93	-	115	5	185	3.75	95 - 185	12.39	19.53	-0.36
		-	115	6	12	2.92	0 - 28	7		0.29
	P97	-	142	6	10	2.67	4 - 24	5.33	-	0.31
	P10	DS	212	8	19.74	3.06	2 - 56	2.67	-	0.23
		RS	219	6	91.92	4.24	88 - 97.3	6.37	-	-0.24
		-	115	3	89	4.36	68 - 92	9.73		0.31
<i>E.u</i>	P93	-	115	4	44	4.66	30 - 59	13.91	36.26	-0.38
		-	115	6	121	3.04	97 - 151	8		-0.37
	P97	-	142	4	51	6.41	13 - 64	13.15	-	-0.47
		DS	212	3	45.22	3.97	15.5 - 65	7.74	12.4	0.39
	P10	DS	212	4	2.27	4.38	0 - 15.5	4.63		-0.31
		RS	219	4	12.91	7.87	7 - 15.5	11.54	-	-0.32

Number of genotypes used for analysis (n), linkage group (LG), position (cM), LOD value, percentage of variance explained by the QTL (PEV,%) and allelic substitution effect (ASE). DS: dry season, RS: rainy season.

Table 5.4: QTLxE interaction for  $\delta^{13}\text{C}$  between the P93 and P97 trials and between the dry season (DS) and the rainy season (RS) in the P10 trial for all markers associated with a QTL.

Parent	Trial	Season	LG	QTL Position	Marker position	Trial or seasonal effect		Marker effect	Interaction effect		
<i>E.g</i>	P93	-	5	185	185.1	0.87		0.002	**	0.03	*
		-	6	12	21	0.81		0.002	**	0.32	
	P97	-	6	10	21	0.81		0.002	**	0.32	
	P10	DS	8	19.74	19.74	7.90E-46	***	0.008	**	0.17	
		RS	6	91.92	91.92	6.30E-46	***	0.002	**	0.79	
<i>E.u</i>	P93	-	3	89	89	0.73		0.005	**	0.05	*
		-	4	44	42	0.86		5.20E-10	***	0.92	
		-	6	121	-	-		-		-	
	P97	-	4	51	42	0.86		5.20E-10	***	0.92	
	P10	DS	3	45.22	45.22	4.10E-47	***	6.50E-05	***	0.003	**
		DS	4	2.27	2.27	2.00E-47	***	1.30E-07	***	0.93	
		RS	4	12.91	12.91	2.10E-47	***	1.80E-07	***	0.46	

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . DS: dry season, RS: rainy season.





parental maps (five for *E.g* and seven for *E.u*) and were distributed on three different LGs in each map (LGs 5, 6 and 8 of *E.g*; LGs 3, 4 and 6 of *E.u*, Table 5.3 and Figure 5.3). The phenotypic variance explained by the QTLs ranged from 2.67% to 12.39% for *E.g*, and from 4.63% to 13.91% for *E.u*. In the parental maps, three colocalizations were found in the different trials: one for *E.g* on LG6 and two for *E.u* on LG3 and LG4. For *E.g*, the colocalization involved QTLs from P93 and P97 explaining 7% and 5.3% of the phenotypic variance. For *E.u*, QTLs from P93 and DS colocated on LG3 and explained 9.7% and 7.7% of the phenotypic variance. A key colocalization in *E.u* was found on LG4 involving highly significant QTLs from all three trials and seasons. The PEV were higher in the P93 (13.9%), P97 (13.2%) and RS (11.5%) than in the DS (4.63%). The QTLs for different compartments of the plants and environments (trials) coincided, suggest a rather stable genetic  $\delta^{13}\text{C}$  architecture.

Among all detected QTLs for  $\delta^{13}\text{C}$  only three presented a significant QTL-by-environment interaction (Table 5.4): one on LG5 (*E.g*) and two on LG3 (*E.u*). As expected, a strong environmental effect (season effect,  $p < 0.001$ ) was found between RS and DS due to the significant differences in climate between the two seasons. Conversely, for a more integrated value of  $\delta^{13}\text{C}$  (P93 and P97), no environmental effect was detected, despite the fact the trials were planted in different years. Overall, most QTLs did not display significant QTL  $\times$  E, which is consistent with the colocalizations found between trials.

### 5.3.3.2 Growth traits

QTLs for growth traits were only searched for in the P10 trial. Over the 20 growth traits analysed, 12 presented at least one QTL (Table S4). A total of 13 and 5 QTLs were detected on two and four LGs in *E.g* and *E.u*, respectively. The percentage of phenotypic variance explained by a QTL ranged from 4.02% (HT10) to 10.12% (D6). As expected, all QTLs detected for *E.g* were grouped in two regions: one on LG3 (eight QTLs) and one on LG6 (five QTLs) given the strong phenotypic correlations between these traits. These two regions were mainly related to diameter and diameter increments even if one QTL for HT10 was also detected on LG3. For *E.u*, the five QTLs detected spread over LG 1, 2, 8 and 9 mainly involving growth increments (three traits out of four). No colocalization was found between growth trait QTLs and  $\delta^{13}\text{C}$  QTLs in P10, in agreement with the low phenotypic correlations observed between these two traits.

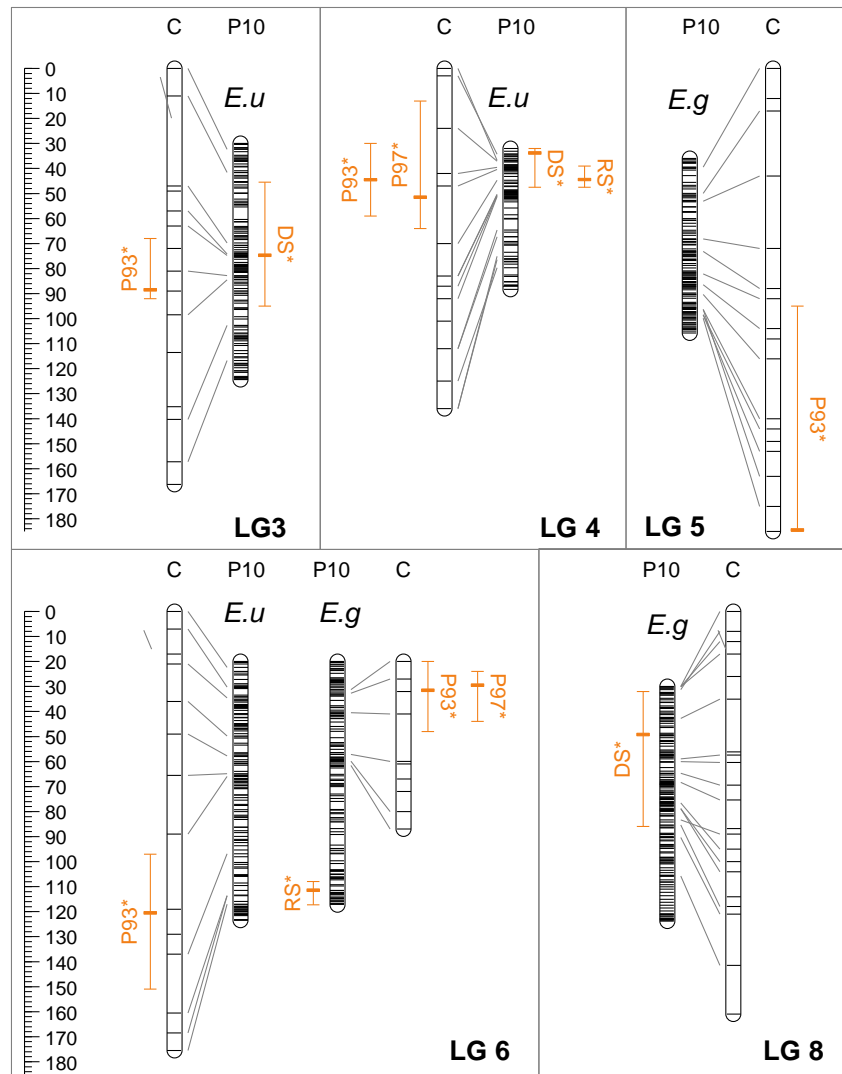


Figure 5.3: Map position of the detected QTLs over five linkage groups (LG) of the *Eucalyptus urophyla* (E.u) and *E. grandis* (E.g) parents.

Table 5.5: Characteristics of the  $\delta^{13}\text{C}$ -QTL regions detected in the P10 trial.

	RS		DS		
	E. u	E. g	E. u	E. g	
LG	4	6	3	4	8
LOD max (cM)	12.91	91.92	45.22	2.27	19.74
Marker name	1151	4531	934	1225	4732
Marker position (bp)	24,565,544	49,540,983	36,055,887	4,721,794	10,939,733
Region:					
upper boundary	23,910,937	49,410,409	35,933,917	4,473,700	10,771,222
lower boundary	25,222,984	49,766,786	37,928,270	5,170,518	12,223,306
size	1,312,047	356,377	1,994,353	696,818	1,452,084
Number of genes:					
total	68	40	55	29	121
with GO terms	41	27	39	25	75
involved in GO enrichment analysis	22	16	9	4	6



### 5.3.4 GO analysis of positional candidate genes

The five  $\delta^{13}\text{C}$ -QTLs detected on four LGs (LG3 and LG4 in *E.u* and LG6 and LG8 in *E.g*) in the P10 trial were screened for GO term enrichment. The physical size of the corresponding regions on the reference genome (Table 5.5) ranged from 356,377 bp (LG6 of *E.g*) to 1,994,353 bp (LG3 of *E.u*). These five QTL regions contained a total of 313 putative genes of which 66% were annotated with GO terms. Regarding the two types of QTLs (dry season DS vs. rainy season RS), the three DS regions totalled 139 annotated genes, while the two RS regions totalled 68 annotated genes. The higher number of genes for the DS region is due to the size of the regions considered: the DS regions being two-fold longer than the RS regions (4,143,255 bp vs. 1,668,424 bp). Out of the 207 annotated genes in the QTL regions, 57 presented a FDR <5% for GO term enrichment compared to the genome (19 in the DS region and 38 in the RS region, listed in Table S5.5). The results of the GO term enrichment for each type of QTL region (DS and RS) are presented in Table S5.6. The DS regions (LG3, 4 and 8) were mainly highlighted by the following GO term components: (i) molecular functions (F): endoribonuclease activity in LG3 and very long-chain fatty acid-CoA ligase activity in LG8, (ii) biological processes (P): unfolded protein response in LG3, regulation of meristem structural organisation in LG8, heat acclimation in LG 3 and 4, and ornithine metabolic process in LG4 and 8. The RS regions (LG4 and 6) were mainly highlighted by the following GO term components: (i) molecular functions (F): sucrose and maltose symporter activities, arbutin and salicin transmembrane transporter activity, and hydroquinone:oxygen oxidoreductase activity in LG6, copper ion binding in LG4 and 6, Rho GDP-dissociation inhibitor activity in LG4, ii/ biological processes (P): glucoside and maltose transport, phloem sucrose loading, pollen germination, lignin catabolic process in LG6, flower development and response to nematodes in LG4 and 6, regulation of stomatal movement, entrainment of circadian clock, negative regulation of photomorphogenesis in LG4.

## 5.4 Discussion

### 5.4.1 The phenotypic variability of $\delta^{13}\text{C}$ is determined by environmental and ontogenic effects

The two levels of integration (season vs. rotation) of  $\delta^{13}\text{C}$  measurements were related to different tree compartments (leaves vs. holocellulose extracted from wood).  $\delta^{13}\text{C}$  measured in



holocellulose was found to be higher on average (up to 2.2‰) than that measured in leaves. Interestingly, a difference of about 2‰ was reported by Bowling *et al.* (2008) between bulk leaf  $\delta^{13}\text{C}$  and  $\delta^{13}\text{C}$  measured in cellulose. Cernusak *et al.* (2009) proposed six hypotheses to explain the enrichment in  $^{13}\text{C}$  observed in non-photosynthetic tissues, such as the difference in biochemical composition between heterotrophic tissues and leaves or the seasonal separation between leaf growth and the growth of heterotrophic tissues.

The increase in WUE with tree age reported by (Donovan & Ehleringer, 1992; Bert *et al.*, 1997; Bond, 2000) could also explain this difference, as wood samples of our study mostly reflected adult compartment of the trees. The increase in WUE with tree age has been linked to different interrelated factors such as the hydraulic limitation due to an increase in tree size (Hubbard *et al.*, 1999; Koch *et al.*, 2004; Forrester *et al.*, 2010), the decline in foliar photosynthetic rates (Yoder *et al.*, 1994; Niinemets, 2002) or the change in foliar structure (Day *et al.*, 2001; Velikova *et al.*, 2008).

The phenotypic variation in  $\delta^{13}\text{C}$  observed in the *E. urophylla* x *E. grandis* full-sib family was large and was related to the period of  $\text{CO}_2$  integration. The minimum range of variation was found in the P93 trial and in the P10-RS (2.7‰) and the maximum was found in the P10-DS (4‰). This is the first report of within family variation in  $\delta^{13}\text{C}$  in *Eucalyptus*. While a slightly higher range of variation was observed at the botanical family level (5.2 ‰ between 65 species of *Eucalyptus* in natural habitats, (Schulze *et al.*, 2006), a smaller range of variation was found between clones of *E. grandis* and *E. grandis* x *camaldulensis* (Bond & Stock, 1990; Le Roux *et al.*, 1996). Such a high level of variability already found between full-sibs of a single family reveals the polygenic nature of carbon isotope discrimination.

In addition to different ranges of variation (higher in the DS than in the RS), this study also highlighted a lack of correlation between the two growing seasons ( $\rho = 0.19$ ). A similar trend was reported in 3-year old *P. elliottii* x *P. caribaea* clones with a correlation of 0.255 between the winter and summer seasons at a dry site, and a much higher correlation (0.9) at a wet site (Prasolova *et al.*, 2003), highlighting the strong impact of environmental factors on genotype ranking, i.e. the presence of a G x E interaction. Likewise, correlations of  $\delta^{13}\text{C}$  between wet and dry treatments were found to vary across years ( $0.34 < \rho < 0.69$ ) in *Vitis vinifera* (Marguerit *et al.*, 2012), while more stable trends were found across years in 20-year old *Picea mariana* trees, with correlations ranging from 0.7 to 0.84 (Johnsen *et al.*, 1999), and



from 0.42 to 0.62 in 3-year old *Castanea sativa* grown in a greenhouse (Casasoli *et al.*, 2004). Marguerit *et al.* (2014) used an experimental trial to estimate genetic correlations between 12-year old *Pinus pinaster* trees planted at three sites with contrasted water availability, and found rather high inter-site correlations (from 0.63 to 0.91). Therefore, our result contrasts with those cited above, most of which reported moderate to strong correlations between years and /or environments. However, the lack of correlation in our study is in agreement with the changes in mature leaf characteristics in *E. urophylla* x *E. grandis* hybrids between the DS (lower SLA) and RS (higher SLA) reported in the Republic of Congo (Nouvellon *et al.*, 2010). We hypothesise that this foliar plasticity (illustrated in Figure S5.1), combined with *E. grandis*' known heteroblasty which displays an abrupt change from juvenile to adult leaf morphology (Zotz *et al.*, 2011), are the main sources of the absence of correlation between the  $\delta^{13}\text{C}$  in the DS and RS .

#### 5.4.2 Genetic architecture of $\delta^{13}\text{C}$ in eucalyptus

This study reports  $\delta^{13}\text{C}$ -QTLs in eucalyptus for the first time. In relation with the moderate heritability found for this trait in other perennials (Johnsen *et al.*, 1999; Prasolova *et al.*, 2001; Lauteri *et al.*, 2004; Brendel *et al.*, 2008), QTLs were detected for all measurements carried out in the three trials (set up in 1993, 1997 and 2010), from two compartments (wood vs. leaves) and two seasons (dry vs. rainy). The genetic architecture matched that reported in other perennials: *Pinus pinaster* (Brendel *et al.*, 2002; Marguerit *et al.*, 2014), *Quercus robur* (Brendel *et al.*, 2008), *Castanea sativa* (Casasoli *et al.*, 2004), *Salix dasyclados* x *viminalis* (Rönnerberg-Wästljung *et al.*, 2005; Weih *et al.*, 2006) and *Vitis vinifera* x *riparia* (Marguerit *et al.*, 2012), i.e. numerous QTLs with small to moderate effects and few QTLs with large effects. This result reflects the numerous physiological processes underlying carbon isotope discrimination (Farquhar *et al.*, 1989) and therefore the multiple regulations involved at the genetic level. The percentage of variance explained by  $\delta^{13}\text{C}$ -QTLs ranged from 2.7% to 13.9% but is likely overestimated especially in the P93 and P97 trials in which 115-142 offspring were phenotyped (Beavis, 1998; Xu, 2003).

The environmental (confounded site and year effects) stability of  $\delta^{13}\text{C}$  at the phenotypic level between the P93 and P97 trials was confirmed at the genetic level by the QTL analysis. Indeed, all  $\delta^{13}\text{C}$ -QTLs detected in the P97 trial colocalized with  $\delta^{13}\text{C}$  QTLs detected in P93.



This stability across the two trials could be explained by the large integration scale of  $\delta^{13}\text{C}$  measured on holocellulose from wood samples representing the duration of the entire rotation, which minimises instantaneous environmental effects. Stable genomic regions were also found between  $\delta^{13}\text{C}$  measured in multiple years in the same trials in oak, grape and pine (Brendel *et al.*, 2008; Marguerit *et al.*, 2012; 2014). Interestingly, a QTL region on LG4 of *E.u* was shared between holocellulose  $\delta^{13}\text{C}$  in trees in the P93 and P97 trials and leaf  $\delta^{13}\text{C}$  in P10-RS. These three QTLs showed the highest effects of all the QTLs detected in *E. urophylla* (see Table 3). However, when the position of the closest markers to the QTL peak was compared with their position on the high density map, the QTL detected for P10-DS on LG4 did not coincide with the P10-RS QTL. As expected, no other colocalization of  $\delta^{13}\text{C}$ -QTL was detected between the DS and RS, due to the small correlation ( $\rho=0.19$ ) between the two seasons. Similar trends were found in *Salix dasyclados* x *S. viminalis* between two watering conditions (Rönnberg-Wästljung *et al.*, 2005). Despite a high phenotypic correlation (0.76) between the two treatments, only one QTL colocalization was found among six  $\delta^{13}\text{C}$ -QTLs. This trend was also reported in *Vitis vinifera* x *riparia* (Marguerit *et al.*, 2012) with no colocalization of  $\delta^{13}\text{C}$  QTLs between control and drought stress conditions.

Altogether these results suggest: (i) the presence of two highly seasonal and/or ontogenic dependent  $\delta^{13}\text{C}$ -QTLs on LG4, which may discriminate DR vs. RS as well as the two successive ontogenic stages of the leaf where heteroblasty is pronounced in eucalyptus (Zotz G *et al.* 2011), and (ii) that the P10-RS  $\delta^{13}\text{C}$ -QTL detected at young stage shares some functional attributes with the  $\delta^{13}\text{C}$ -QTL integrated over time. Although carbon discrimination over the course of the RS appeared to be an important contributor to carbon isotope composition in the holocellulose of wood, because our study was limited in time (1<sup>st</sup> DS and 2<sup>nd</sup> RS), we cannot rule out the possibility that DS  $\delta^{13}\text{C}$ -QTLs may also play an important role at a later stage. This hypothesis remains to be tested.

#### 5.4.3 Genes underlying $\delta^{13}\text{C}$ -QTLs support differences in leaf $\delta^{13}\text{C}$ between the dry and rainy seasons

Providing that high resolution linkage maps (combining high marker density with high accuracy of genetic distances) are anchored on an annotated reference genome sequence, it is possible to explore positional candidate genes underlying QTLs, as illustrated by Monclus *et al.* (2012) in *Populus* spp. In the present study, we used this strategy to characterise leaf  $\delta^{13}\text{C}$ -



QTLs, in the DS and RS taking advantage of the high resolution eucalyptus genetic maps based on 1,025 full-sibs genotyped with a 6k SNP array (Bartholomé *et al.*, 2014). Considering that genomic regions surrounding LOD peaks (*ca* 1cM) should contain the underlying gene (Price, 2006), GO term enrichment analyses were performed based on the regions defined by the two closest markers to the marker associated with the LOD peak. Our results revealed distinct patterns of enrichment according to the two QTL categories (DS vs. RS).

The three  $\delta^{13}\text{C}$ -DS QTL regions contained 19 genes with main roles in stress response based on their TAIR annotation. The response to stresses is brought by GO terms for the three LGs harbouring  $\delta^{13}\text{C}$ -DS QTL (GO:0010286 for LG3 and LG4 and GO:0006591 for LG4 and LG8), with two genes corresponding to the F-box gene family for which differential expression has been reported during heat acclimation in *At* suspension-culture cells (Chan *et al.* 2006), and with two others involved in the ornithine metabolic process, which has been suggested to modulate tolerance to abiotic stresses in *Arabidopsis* (Kalamaki *et al.* 2009). The QTL localized on LG8 is enriched in two other GO terms (GO:0007172, GO:0009934) brought by two CLV1 genes encoding a Leucine-rich repeat receptor kinase (LRR-RKs), known to control shoot and floral meristem size, and to establish and maintain floral meristem identity. In *Medicago trunculata*, de Lorenzo *et al.* (2009) reported the identification of an LRR-RK gene (*Srlk*) in the regulation of the adaptation of roots to salt stress. The last interesting GO term of LG8 is related to very long chain fatty acid-CoA ligase activity (GO:0031957). The corresponding two members of the AMP-dependent synthetase and ligase family protein (*CER8*) are involved in cuticular waxes and cutin biosynthesis for the maintenance of plant water status ((Lü *et al.*, 2012; Li-Beisson *et al.*, 2013). Interestingly, the cuticular phenotypes of *cer9* (a mutant of *ECERIFERUM9* gene in *Arabidopsis*) are associated with delayed onset of wilting in plants experiencing water deficit, lower transpiration rates, and improved water use efficiency (WUE) measured as carbon isotope discrimination. Altogether, these results suggest that the variability of  $\delta^{13}\text{C}$  in leaves of 10-month old eucalyptus in the DS is related to a response to abiotic stress and more particularly to water deficit. Marguerit *et al.* (2012) also reported positional candidate genes involved in stress response colocalized with leaf  $\delta^{13}\text{C}$ -QTLs in *Vitis vinifera* x *riparia* under water constraint.





Conversely, the 38 genes of the  $\delta^{13}\text{C}$ -RS QTLs (in LG 4 and 6) were enriched in a more variable set of GO terms. For these  $\delta^{13}\text{C}$  QTLs, it was more difficult to highlight a single process to characterise the genetic control of such a trait as illustrated by seven genes corresponding to one member of the *Arabidopsis thaliana* laccases, a gene family known to be involved in different developmental stages in different organs (Turlapati *et al.*, 2011), two genes (SEC6 and ATSERK1) either involved in cytokinesis (Wu *et al.*, 2013) or known to be regulated by auxin during morphogenesis (Elhiti *et al.*, 2013), a KNOX/ELK homeobox transcription factor involved in meristem and organ development (Liu *et al.*, 2013), and two genes (SRF8 and EBS) related to the timing of flowering initiation in *Arabidopsis* (Piñeiro *et al.*, 2003). However, more specific and interesting GO terms are provided by the QTL on LG4. This region is enriched with five genes specifically linked to stomatal functioning: two genes (ATCOP1) encoding Transducin/WD40 repeat-like proteins linked to photomorphogenesis and the circadian clock (Pacín *et al.*, 2013; Zheng *et al.*, 2013), two processes indirectly involved in stomatal regulation (as reviewed (Araújo *et al.*, 2011), two encoding Immunoglobulin E-set proteins (SCN1 gene) and one encoding protein kinase (HT1), which are linked to stomatal closure/opening in *Arabidopsis* (Jeon *et al.*, 2008; Jung *et al.*, 2008). The discovery of genes involved in the stomatal functioning as putative positional candidate genes for RS- $\delta^{13}\text{C}$  is probably the most relevant result of the GO terms enrichment analysis, stressing that this QTL is involved in the variability of stomatal functioning, i.e. in the first step of  $\text{CO}_2$  assimilation.

To conclude, both the genetic architecture and differences in the functional characterisation of the two  $\delta^{13}\text{C}$ -QTL categories (RS vs. DS) highlight the polygenic and multifunctional nature of leaf  $\delta^{13}\text{C}$ . These different mechanisms certainly play a role in intrinsic water use efficiency, but their relative contribution to plant transpiration efficiency remains to be investigated.

#### 5.4.4 Relation between $\delta^{13}\text{C}$ , WUE and growth

The variability among genotypes for  $\delta^{13}\text{C}$  was found to be independent of growth whatever the level of integration of  $\delta^{13}\text{C}$  over the season or throughout the rotation. Indeed, almost no significant correlation between  $\delta^{13}\text{C}$  and cumulative growth was found in the three trials except between the  $\delta^{13}\text{C}$  P10-DS. The absence of a correlation between these two traits



was also reported in *Larix occidentalis* (Zhang *et al.*, 1994), *Araucaria cunninghamii* (Xu *et al.*, 2003), *Populus deltoides* x *P. nigra* (Monclus *et al.*, 2005; Monclus *et al.*, 2006). However no clear trend of positive, null, or negative correlation between growth and  $\delta^{13}\text{C}$  is cited in the literature for perennial species (Zhang *et al.*, 1994; Nguyen-Queyrens *et al.*, 1998; Johnsen *et al.*, 1999; Prasolova *et al.*, 2001; Brendel *et al.*, 2002; Xu *et al.*, 2003; Weih *et al.*, 2006; Monclus *et al.*, 2012). The most likely independence between the two traits in eucalyptus is reflected in the absence of QTL colocalization in our study as is the case in numerous tree species: *Pinus pinaster*, *Quercus robur* (Brendel *et al.*, 2008), *Castanea sativa* (Casasoli *et al.*, 2004). Using an association genetics approach in *Pinus taeda*, Cumbie *et al.* (2011) did not find any SNP associated with both  $\delta^{13}\text{C}$  and growth either. Given that  $\delta^{13}\text{C}$  could be considered as a good proxy for WUE (as suggested by our study in the rainy season and by others), we hypothesise that the genetic factors underlying the variability of WUE and growth are decoupled. Interestingly, the low correlations between P10-DS leaf  $\delta^{13}\text{C}$  and growth at young stages, reinforce the hypothesis that foliar  $\delta^{13}\text{C}$  integrated over the DS is a signature of s response to water stress rather than WUE *per se*.

In forest tree breeding programmes mainly focused on growth traits, the possible absence of genetic correlation between growth traits and  $\delta^{13}\text{C}$  and by extension WUE, could be advantageous in the context of climate change (Dillen *et al.*, 2008; Marguerit *et al.*, 2014). Indeed, it may be possible to improve the WUE of forest tree varieties whose growth ability has already been improved.

### Authors' contributions

AM, JB were responsible for growth measurements and field sampling. DB, JB for preparation of the samples. JB for genetic linkage analysis and QTL analysis. JMG, BS for GO analysis. JB, JMG, CP, OB, DB wrote the manuscript. JMG: Funding and overall supervision. All the authors read and approved the final manuscript.

### Acknowledgments

The authors thank the CRDPI (*Centre de Recherches sur la Durabilité des Plantations Industrielles*) in the Republic of Congo for setting up the trials. The authors also thank Régis Burlett and Alexandre Bosc for their help with climatic data mining. Isotopic measurements were performed by the Technical Platform of Functional Ecology (OC 081) at INRA (Nancy, France). This study was supported by grants from FEDER (ABIOTEN project, n° Presage 32973), ERANET Plant KBBE (ANR(FR)-10-KBBE-0007 “34Joule”) and ANR (Project WUETREE). JB received a PhD fellowship from the BIOS department of CIRAD.



## Supplemental figures and tables



Figure S5.1: Illustration of heteroblasty (different leaf forms depending on age) for one single genotype (tree #502). In *Eucalyptus* hybrids grown in the Republic of Congo, heteroblasty is combined with a seasonal effect (dry vs. rainy seasons). Left: 10-month old fully extended leaves sampled at the end of the dry season. Right: 18-month old fully extended leaves sampled at the end of the rainy season.

Table S5.1: Characteristics of the genetic maps for both parents (*E. urophylla* and *E. grandis*). The components maps (P93 and P97) used to build the consensus maps are also indicated.

	E.u			E.g			E.u	E.g
	P93 <sup>a</sup>	P97 <sup>a</sup>	Consensus <sup>b</sup>	P93 <sup>a</sup>	P97 <sup>a</sup>	Consensus <sup>b</sup>	P10 <sup>c</sup>	P10 <sup>c</sup>
Number of LGs	11	13	11	11	11	11	11	11
Map length (cM)	1502	1179	1609	1270	1021	1370	885.9	821.7
Average LG length (cM)	137	91	146.3	115	93	124.5	80.5	74.7
Number of framework markers	146	106	180	118	86	147	1773	1832
Mean distance between markers	11	15	10.3	12	15	9.8	0.65	0.58

<sup>a</sup>: from Bartholomé et al. 2013

<sup>b</sup>: this study

<sup>c</sup>: from Bartholomé et al. 2014



Table S5.2: Summary of growth traits for P10. The following features are indicated: Number of values (N), minimum (Min), maximum (Max), mean, standard deviation (SD) and coefficient of phenotypic variation (CVp). HT: height, D: diameter, RS: rainy season, DS: dry season

Traits	Season	N	Min	Max	Mean	SD	CVp
HT3	-	220	0.18	0.9	0.58	0.13	0.23
HT10	-	220	0.6	5.5	2.92	0.85	0.29
HT18	-	216	5	12.3	9.09	1.34	0.15
HT23	-	220	6.1	14.5	11.14	1.47	0.13
HT3_10	-	220	0.04	0.66	0.32	0.11	0.34
HT10_18	RS	216	0.36	0.9	0.71	0.09	0.13
HT18_23	DS	216	0.19	0.83	0.5	0.11	0.22
D3	-	214	1.2	11.7	5.54	1.95	0.35
D6	-	218	3.9	38.9	20.62	6.4	0.31
D10	-	220	7	78	41.24	13.19	0.32
D15	-	220	22	116	76.34	16.1	0.21
D18	-	219	54	148	99.27	19.14	0.19
D23	-	219	55.5	171.5	116.8	21.11	0.18
D3_6	RS	212	0.9	10.27	5.03	1.72	0.34
D3_10	-	214	0.79	9.11	4.91	1.68	0.34
D6_10	DS	218	0.54	8.94	4.57	1.85	0.41
D10_15	RS	220	2.88	9.67	6.75	1.11	0.16
D10_18	RS	220	3.52	9.39	6.64	1.13	0.17
D15_18	RS	218	1.3	10.03	6.1	1.51	0.25
D18_23	DS	218	1	7.57	4.26	0.99	0.23
$\delta^{13}\text{C}$	DS	212	-31.5	-27.5	-29.89	0.71	-
$\delta^{13}\text{C}$	RS	219	-30.2	-27.5	-28.97	0.47	-



Table S5.3: Spearman correlation coefficient between  $\delta^{13}\text{C}$  and growth traits for all trials: P93, P97 and P10. Significant correlations are indicated by \*. Height: HT, Diameter: D, Circumference: C

Traits	P93	Traits	P97	Traits	P97	Traits	P10_DS	P10_RS
HT14	0.18	HT7	0.07	C7	0.02	HT3	0.13	-0.01
HT26	0.17	HT8	0.06	C8	0.08	HT10	0.3 *	0.12
HT39	0.17	HT9	0.01	C9	0.08	HT18	0.21 *	0.05
HT51	0.11	HT10	-0.01	C10	0.11	HT22	0.19 *	0.04
HT59	0.08	HT11	0.05	C11	0.13	HT3_10	0.29 *	0.12
HT14_26	0.04	HT12	0.08	C12	0.13	HT10_18	0.05	-0.03
HT26_39	0	HT13	0.07	C13	0.14	HT18_22	-0.05	-0.04
HT39_51	-0.09	HT15	0.16	C15	0.16	D3	0.23 *	0.05
HT51_59	-0.07	HT18	0.16	C18	0.17 *	D6	0.28 *	0.05
C14	-0.02	HT21	0.19 *	C21	0.16	D10	0.29 *	0.1
C26	0.01	HT24	0.15	C24	0.14	D15	0.24 *	0.11
C39	-0.03	HT27	0.11	C27	0.13	D18	0.19 *	0.09
C51	-0.05	HT30	0.09	C30	0.13	D22	0.18 *	0.1
Cir59	-0.1	HT34	0.08	C34	0.1	D3_6	0.26 *	0.04
C14_26	-0.04	HT44	0.1	C44	0.11	D3_10	0.28 *	0.09
C26_39	-0.1	HT52	0.11	C52	0.13	D6_10	0.23 *	0.11
C39_51	-0.13	HT62	0.09	C62	0.13	D10_15	-0.07	0.1
C51_59	-0.28 *	HT7_8	0.01	C7_8	0.07	D10_18	-0.01	0.07
		HT8_9	-0.05	C8_9	0.02	D15_18	0.06	0.02
		HT9_10	-0.03	C9_10	0.08	D18_22	-0.04	0.02
		HT10_11	0.16	C10_11	0.09	$\delta^{13}\text{C}_{\text{DS}}$		0.19 *
		HT11_12	0.15	C11_12	0.05	$\delta^{13}\text{C}_{\text{RS}}$	0.19 *	
		HT12_13	0.07	C12_13	0.1			
		HT13_15	0.2 *	C13_15	0.11			
		HT15_18	0.15	C15_18	0.15			
		HT18_21	0.12	C18_21	0.06			
		HT21_24	-0.08	C21_24	0.06			
		HT24_27	-0.14	C24_27	0			
		HT27_30	0.03	C27_30	0.09			
		HT30_34	0.13	C30_34	0.03			
		HT34_44	0.1	C34_44	0.12			
		HT44_52	0.12	C44_52	0.13			
		HT52_62	0.05	C52_62	-0.02			



Table S5.4: QTLs for growth traits detected by composite interval mapping in both parents (E.g and E.u) in the P10 trial. The following features are shown: number of genotypes used for analysis (n), linkage group (LG), position (cM), LOD value (significant at 5% genome-wide level), percentage of variance explained by the QTL (PEV, %) the allelic substitution effect (ASE).

Parental map	Trait	n	LG	Position	LOD	BCI	PEV	PEV tot	ASE
E.g	HT10	220	3	45.78	3.27	32 - 53.5	4.02	-	-0.34
	D6	218	3	45.78	5.31	37.5 - 48.6	10.12	12.43	-4.16
	D6	218	6	75.04	2.87	32 - 88.5	4.47	12.43	2.76
	D10	220	3	45.78	3.77	35.3 - 53	7.26	-	-7.27
	D15	220	3	46.5	3.51	32.5 - 53	6.59	11.13	-8.48
	D15	220	6	75.04	3.29	66.5 - 85.5	6.71	11.13	8.51
	D18	219	3	46.5	3.98	35 - 53	6.82	-	-10.28
	D22	219	3	46.5	4.45	34.5 - 53	6.75	10.57	-11.24
	D22	219	6	75.04	3.86	66.5 - 83.7	5.85	10.57	10.41
	D3_6	212	3	45.78	5.3	41.3 - 49	9.86	13.64	-1.1
	D3_6	212	6	75.04	3.69	68 - 87	6.06	13.64	0.86
	D3_10	214	3	46.5	3.58	31.5 - 53	6.62	10.63	-0.88
	D3_10	214	6	75.04	2.99	55 - 96	5.86	10.63	0.82
E.u	HT18	216	9	1.77	3.04	1 - 15	6.22	-	-0.67
	HT10_18	216	1	11.74	3.43	7 - 27.1	4.81	12.71	-0.04
	HT10_18	216	9	2	4.69	0 - 12	8.49	12.71	-0.05
	D10_15	220	2	102.5	3.38	74.9 - 105.5	5.54	-	0.53
	D18_22	218	8	51.6	4.11	37 - 77	7.56	-	0.55

HT: height, D: diameter


Table S5.5: List of genes underlying  $\delta^{13}\text{C}$  QTL region for which GO annotation present an enrichment.

QTL	Parent	LG	Gene ID	AT name	e-value	Bit score	AT symbol	Function
DS	<i>E.u</i>	3	Eucgr.C02014.1	AT5G10695.1	9.00E-15	61.6		unknown protein
DS	<i>E.u</i>	3	Eucgr.C02015.1	AT5G10695.1	6.00E-20	74.7		unknown protein
DS	<i>E.u</i>	3	Eucgr.C02016.1	AT5G10695.1	2.00E-31	103		unknown protein
DS	<i>E.u</i>	3	Eucgr.C02029.1	AT2G17520.1	2.00E-132	405	ATIRE1-2, IRE1-2, IRE1A	Endoribonuclease/protein kinase IRE1-like
DS	<i>E.u</i>	3	Eucgr.C02031.1	AT2G17520.1	4.00E-110	345	ATIRE1-2, IRE1-2, IRE1A	Endoribonuclease/protein kinase IRE1-like
DS	<i>E.u</i>	3	Eucgr.C02037.1	AT2G17520.1	4.00E-35	131	ATIRE1-2, IRE1-2, IRE1A	Endoribonuclease/protein kinase IRE1-like
DS	<i>E.u</i>	3	Eucgr.C02042.1	AT2G17520.1	8.00E-138	421	ATIRE1-2, IRE1-2, IRE1A	Endoribonuclease/protein kinase IRE1-like
DS	<i>E.u</i>	3	Eucgr.C02048.1	AT2G17520.1	2.00E-137	419	ATIRE1-2, IRE1-2, IRE1A	Endoribonuclease/protein kinase IRE1-like
DS	<i>E.u</i>	3	Eucgr.C02083.1	AT5G24360.1	3.00E-45	161	ATIRE1-1, IRE1-1	inositol requiring 1-1
DS	<i>E.u</i>	4	Eucgr.D00282.1	AT2G43190.2	5.00E-84	254		ribonuclease P family protein
DS	<i>E.u</i>	4	Eucgr.D00284.1	AT5G46170.1	1.00E-177	500		F-box family protein
DS	<i>E.u</i>	4	Eucgr.D00286.1	AT5G46170.1	0.0	518		F-box family protein
DS	<i>E.u</i>	4	Eucgr.D00287.1	AT5G46180.1	0.0	713	DELTA-OAT	ornithine-delta-aminotransferase
DS	<i>E.g</i>	8	Eucgr.H00899.1	AT4G01900.1	7.00E-54	169	GLB1, PII	GLNB1 homolog
DS	<i>E.g</i>	8	Eucgr.H00913.1	AT2G47240.1	0.0	971	CER8, LACS1	AMP-dependent synthetase and ligase family protein
DS	<i>E.g</i>	8	Eucgr.H00914.1	AT2G47240.1	0.0	934	CER8, LACS1	AMP-dependent synthetase and ligase family protein
DS	<i>E.g</i>	8	Eucgr.H00963.1	AT1G75820.1	0.0	1184	ATCLV1, CLV1, FAS3, FLO5	Leucine-rich receptor-like protein kinase family protein
DS	<i>E.g</i>	8	Eucgr.H00964.1	AT1G75820.1	0.0	1148	ATCLV1, CLV1, FAS3, FLO5	Leucine-rich receptor-like protein kinase family protein
DS	<i>E.g</i>	8	Eucgr.H00972.1	AT4G24190.1	0.0	721	AtHsp90-7, AtHsp90.7, HSP90.7, SHD	Chaperone protein htpG family protein
RS	<i>E.u</i>	4	Eucgr.D01339.1	AT1G11820.2	0.0	601		O-Glycosyl hydrolases family 17 protein
RS	<i>E.u</i>	4	Eucgr.D01340.1	AT3G56190.1	3.00E-67	208	ALPHA-SNAP2, ASNAP	alpha-soluble NSF attachment protein 2
RS	<i>E.u</i>	4	Eucgr.D01341.1	AT3G57610.1	0.0	778	ADSS	adenylosuccinate synthase
RS	<i>E.u</i>	4	Eucgr.D01344.1	AT4G22010.1	0.0	870	sks4	SKU5 similar 4
RS	<i>E.u</i>	4	Eucgr.D01347.1	AT4G28200.1	0.0	700		
RS	<i>E.u</i>	4	Eucgr.D01351.1	AT1G11910.1	0.0	734	APA1, ATAPA1	aspartic proteinase A1
RS	<i>E.u</i>	4	Eucgr.D01352.1	AT1G11920.1	0.0	560		Pectin lyase-like superfamily protein
RS	<i>E.u</i>	4	Eucgr.D01358.1	AT4G22120.1	0.0	1177		ERD (early-responsive to dehydration stress) family protein
RS	<i>E.u</i>	4	Eucgr.D01360.1	AT4G22130.1	0.0	913	SRF8	STRUBBELIG-receptor family 8
RS	<i>E.u</i>	4	Eucgr.D01362.1	AT4G22140.2	3.00E-146	406	EBS	PHD finger family protein / bromo-adjacent homology (BAH) domain-containing protein
RS	<i>E.u</i>	4	Eucgr.D01364.1	AT1G62360.1	1.00E-136	394	BUM, BUM1, SHL, STM, WAM, WAM1	KNOX/ELK homeobox transcription factor
RS	<i>E.u</i>	4	Eucgr.D01370.1	AT2G32950.1	1.00E-101	309	ATCOP1, COP1,	Transducin/WD40 repeat-like superfamily protein





RS	<i>E.u</i>	4	Eucgr.D01371.1	AT2G37050.3	0.0	1153	DET340, EMB168, FUS1	Leucine-rich repeat protein kinase family protein
RS	<i>E.u</i>	4	Eucgr.D01376.1	AT2G32950.1	3.00E-141	422	ATCOP1, COP1, DET340, EMB168, FUS1	Transducin/WD40 repeat-like superfamily protein
RS	<i>E.u</i>	4	Eucgr.D01378.1	AT2G37050.1	5.00E-177	516		Leucine-rich repeat protein kinase family protein
RS	<i>E.u</i>	4	Eucgr.D01381.1	AT4G05520.1	0.0	911	ATEHD2, EHD2	EPS15 homology domain 2
RS	<i>E.u</i>	4	Eucgr.D01382.1	AT4G05520.2	0.0	674	ATEHD2, EHD2	EPS15 homology domain 2
RS	<i>E.u</i>	4	Eucgr.D01384.1	AT1G62400.1	0.0	621	HT1	Protein kinase superfamily protein
RS	<i>E.u</i>	4	Eucgr.D01394.1	AT2G32300.1	2.00E-20	85.5	UCC1	uclacyanin 1
RS	<i>E.u</i>	4	Eucgr.D01404.1	AT3G07880.1	1.00E-47	152	SCN1	Immunoglobulin E-set superfamily protein
RS	<i>E.u</i>	4	Eucgr.D01405.1	AT5G08335.1	4.00E-22	85.9	ATICMTB, ATSTE14B, ICMTB	Isoprenylcysteine carboxyl methyltransferase (ICMT) family
RS	<i>E.u</i>	4	Eucgr.D01410.1	AT3G07880.1	2.00E-111	320	SCN1	Immunoglobulin E-set superfamily protein
RS	<i>E.g</i>	6	Eucgr.F04134.1	AT1G34430.1	7.00E-109	318	EMB3003	2-oxoacid dehydrogenases acyltransferase family protein
RS	<i>E.g</i>	6	Eucgr.F04137.1	AT1G22710.1	0.0	555	ATSUC2, SUC2, SUT1	sucrose-proton symporter 2
RS	<i>E.g</i>	6	Eucgr.F04138.1	AT1G22710.1	0.0	577	ATSUC2, SUC2, SUT1	sucrose-proton symporter 2
RS	<i>E.g</i>	6	Eucgr.F04144.1	AT1G22710.1	0.0	605	ATSUC2, SUC2, SUT1	sucrose-proton symporter 2
RS	<i>E.g</i>	6	Eucgr.F04145.1	AT1G71860.1	3.00E-159	449	ATPTP1, PTP1	protein tyrosine phosphatase 1
RS	<i>E.g</i>	6	Eucgr.F04146.1	AT1G34300.1	0.0	707		lectin protein kinase family protein
RS	<i>E.g</i>	6	Eucgr.F04151.1	AT1G71830.1	0.0	997	ATSERK1, SERK1	somatic embryogenesis receptor-like kinase 1
RS	<i>E.g</i>	6	Eucgr.F04152.1	AT1G34300.1	0.0	882		lectin protein kinase family protein
RS	<i>E.g</i>	6	Eucgr.F04154.1	AT5G09360.1	0.0	575	LAC14	laccase 14
RS	<i>E.g</i>	6	Eucgr.F04156.1	AT1G71820.1	0.0	1294	SEC6	SEC6
RS	<i>E.g</i>	6	Eucgr.F04157.1	AT5G09360.1	4.00E-113	342	LAC14	laccase 14
RS	<i>E.g</i>	6	Eucgr.F04158.1	AT5G09360.1	1.00E-125	374	LAC14	laccase 14
RS	<i>E.g</i>	6	Eucgr.F04159.1	AT5G09360.1	2.00E-38	135	LAC14	laccase 14
RS	<i>E.g</i>	6	Eucgr.F04160.1	AT5G09360.1	0.0	536	LAC14	laccase 14
RS	<i>E.g</i>	6	Eucgr.F04162.1	AT5G09360.1	9.00E-170	491	LAC14	laccase 14
RS	<i>E.g</i>	6	Eucgr.F04163.1	AT5G09360.1	0.0	561	LAC14	laccase 14


Table S5.6: Enrichment analysis in GO terms for genes underlying  $\delta^{13}\text{C}$  QTL region in the dry and rainy seasons

QTL	GO Term	Name of the most specific terms	Type	FDR	A	B	C	D	LG	Genes ID
DS	GO:0006987	activation of signaling protein activity involved in unfolded protein response	P	1.20E-06	6	13	126	24854	3	Eucgr.C02029.1, Eucgr.C02048.1, Eucgr.C02037.1, Eucgr.C02042.1, Eucgr.C02031.1, Eucgr.C02083.1
	GO:0004521	endoribonuclease activity	F	5.30E-05	7	68	125	24799	3, 4	Eucgr.C02029.1, Eucgr.C02048.1, Eucgr.C02037.1, Eucgr.C02042.1, Eucgr.C02031.1, Eucgr.C02083.1, Eucgr.D00282.1
	GO:0042406	extrinsic to endoplasmic reticulum membrane	C	2.00E-06	6	19	126	24848	3	Eucgr.C02029.1, Eucgr.C02048.1, Eucgr.C02037.1, Eucgr.C02042.1, Eucgr.C02031.1, Eucgr.C02083.1
	GO:0007172	signal complex assembly	P	1.30E-02	2	1	130	24866	8	Eucgr.H00964.1, Eucgr.H00963.1
	GO:0009934	regulation of meristem structural organization	P	1.40E-02	3	14	129	24853	8	Eucgr.H00964.1, Eucgr.H00963.1, Eucgr.H00972.1
	GO:0031957	very long-chain fatty acid-CoA ligase activity	F	4.80E-03	2	0	130	24867	8	Eucgr.H00914.1, Eucgr.H00913.1
	GO:0010286	heat acclimation	P	1.80E-02	5	86	127	24781	3, 4	Eucgr.C02016.1, Eucgr.C02015.1, Eucgr.C02014.1, Eucgr.D00286.1, Eucgr.D00284.1
	GO:0006591	ornithine metabolic process	P	3.70E-02	2	3	130	24864	4, 8	Eucgr.D00287.1, Eucgr.H00899.1
	GO:0042946	glucoside transport	P	3.40E-03	2	0	64	24933	6	Eucgr.F04138.1, Eucgr.F04137.1
	GO:0046274	lignin catabolic process	P	5.30E-03	4	51	62	24882	6	Eucgr.F04160.1, Eucgr.F04154.1, Eucgr.F04163.1, Eucgr.F04162.1
RS	GO:0009915	phloem sucrose loading	P	6.10E-04	3	2	63	24931	6	Eucgr.F04144.1, Eucgr.F04138.1, Eucgr.F04137.1
	GO:0009846	pollen germination	P	3.40E-03	4	33	62	24900	6	Eucgr.F04144.1, Eucgr.F04138.1, Eucgr.F04156.1, Eucgr.F04137.1
	GO:0015768	maltose transport	P	3.40E-03	2	0	64	24933	6	Eucgr.F04138.1, Eucgr.F04137.1
	GO:0008506	sucrose:hydrogen symporter activity	F	6.10E-04	3	2	63	24931	6	Eucgr.F04144.1, Eucgr.F04138.1, Eucgr.F04137.1
	GO:0005364	maltose:hydrogen symporter activity	F	3.40E-03	2	0	64	24933	6	Eucgr.F04138.1, Eucgr.F04137.1



QTL	GO Term	Name of the most specific terms	Type	FDR	A	B	C	D	LG	Genes ID
	GO:0042951	arbutin transmembrane transporter activity	F	3.40E-03	2	0	64	24933	6	Eucgr.F04138.1, Eucgr.F04137.1
	GO:0042950	salicin transmembrane transporter activity	F	3.40E-03	2	0	64	24933	6	Eucgr.F04138.1, Eucgr.F04137.1
	GO:0052716	hydroquinone:oxygen oxidoreductase activity	F	5.30E-03	4	51	62	24882	6	Eucgr.F04160.1, Eucgr.F04154.1, Eucgr.F04163.1, Eucgr.F04162.1
	GO:0009624	response to nematode	P	2.40E-02	4	97	62	24836	4, 6	Eucgr.D01360.1, Eucgr.F04144.1, Eucgr.F04138.1, Eucgr.F04137.1 Eucgr.D01347.1, Eucgr.D01364.1, Eucgr.D01405.1, Eucgr.D01344.1, Eucgr.D01362.1, Eucgr.D01360.1, Eucgr.F04138.1, Eucgr.F04137.1, Eucgr.F04151.1
	GO:0009908	flower development	P	3.10E-02	9	775	57	24158	4, 6	Eucgr.D01344.1, Eucgr.D01394.1, Eucgr.F04159.1, Eucgr.F04163.1, Eucgr.F04158.1, Eucgr.F04162.1, Eucgr.F04157.1, Eucgr.F04160.1, Eucgr.F04154.1,
	GO:0005507	copper ion binding	F	1.40E-03	9	364	57	24569	4, 6	Eucgr.D01410.1, Eucgr.D01404.1
RS	GO:0005094	Rho GDP- dissociation inhibitor activity	F	9.00E-03	2	2	64	24931	4	Eucgr.D01371.1, Eucgr.D01351.1, Eucgr.D01378.1, Eucgr.D01382.1, Eucgr.D01381.1,
	GO:0009506	plasmodesma	C	6.60E-03	11	895	55	24038	4, 6	Eucgr.D01344.1, Eucgr.D01340.1, Eucgr.F04144.1, Eucgr.F04138.1, Eucgr.F04156.1, Eucgr.F04137.1
	GO:0005576	extracellular region	C	2.40E-02	13	1518	53	23415	4, 6	Eucgr.D01371.1, Eucgr.D01352.1, Eucgr.D01351.1, Eucgr.D01378.1, Eucgr.D01364.1, Eucgr.D01344.1, Eucgr.D01339.1, Eucgr.D01360.1, Eucgr.D01341.1,



QTL	GO Term	Name of the most specific terms	Type	FDR	A	B	C	D	LG	Genes ID
RS	GO:0005886	plasma membrane	C	4.50E-02	18	2846	48	22087	4, 6	Eucgr.F04163.1, Eucgr.F04162.1, Eucgr.F04160.1, Eucgr.F04154.1, Eucgr.D01371.1, Eucgr.D01378.1, Eucgr.F04134.1, Eucgr.D01364.1, Eucgr.F04152.1, Eucgr.D01382.1, Eucgr.D01358.1, Eucgr.F04151.1, Eucgr.D01381.1, Eucgr.F04146.1, Eucgr.D01339.1, Eucgr.F04145.1, Eucgr.F04144.1, Eucgr.D01360.1, Eucgr.F04138.1, Eucgr.F04156.1, Eucgr.D01340.1, Eucgr.F04137.1
										Eucgr.D01384.1, Eucgr.D01370.1, Eucgr.D01376.1
										Eucgr.D01370.1, Eucgr.D01376.1
										Eucgr.D01370.1, Eucgr.D01376.1
										Eucgr.D01370.1, Eucgr.D01376.1
	GO:0010119	regulation of stomatal movement	P	4.10E-02	3	45	63	24888	4	
	GO:0009649	entrainment of circadian clock	P	4.30E-02	2	8	64	24925	4	
	GO:0010100	negative regulation of photomorphogenesis	P	4.30E-02	2	8	64	24925	4	



## References

- Allen CD, Macalady AK, Chenchouni H, Bachelet D, McDowell N, Vennetier M, Kitzberger T, Rigling A, Breshears DD, Hogg EH, *et al.* 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management* **259**(4): 660-684.
- Almeida AC, Soares JV, Landsberg JJ, Rezende GD. 2007. Growth and water balance of *Eucalyptus grandis* hybrid plantations in Brazil during a rotation for pulp production. *Forest Ecology and Management* **251**(1-2): 10-21.
- Araújo WL, Fernie AR, Nunes-Nesi A. 2011. Control of stomatal aperture: A renaissance of the old guard. *Plant Signaling & Behavior* **6**(9): 1305-1311.
- Bartholomé J, Mandrou E, Mabiala A, Nabihoudine I, Klopp C, Plomion C, Gion J-M. 2014. High-resolution genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* improve BRASUZ1 reference genome assembly and reveal segregation distortion regions. *In prep.*
- Bartholomé J, Salmon F, Vigneron P, Bouvet J-M, Plomion C, Gion J-M. 2013. Plasticity of primary and secondary growth dynamics in *Eucalyptus* hybrids: a quantitative genetics and QTL mapping perspective. *BMC Plant Biology* **13**(1): 120.
- Battaglia M, Bruce J, Brack C, Baker T 2009. Climate Change and Australia's plantation estate: Analysis of vulnerability and preliminary investigation of adaptation options. Canberra: Forest & Wood Products Australia. 125.
- Beavis WD. 1998. QTL analyses: power, precision, and accuracy. *Molecular dissection of complex traits* **1998**: 145-162.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate - A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological* **57**(1): 289-300.
- Bert D, Leavitt SW, Dupouey J-L. 1997. Variations of wood  $\delta^{13}\text{C}$  and water-use efficiency of *Abies alba* during the last century *Ecology* **78**(5): 1588-1596.
- Bond BJ. 2000. Age-related changes in photosynthesis of woody plants. *TRENDS in Plant Science* **5**(8): 349-353.
- Bond WJ, Stock WD. 1990. Preliminary Assessment of the Grading of *Eucalyptus* Clones Using Carbon Isotope Discrimination. *South African Forestry Journal* **154**(1): 51-55.
- Bowling DR, Pataki DE, Randerson JT. 2008. Carbon isotopes in terrestrial ecosystem pools and CO<sub>2</sub> fluxes. *New Phytologist* **178**(1): 24-40.
- Breda N, Huc R, Granier A, Dreyer E. 2006. Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences. *Annals of Forest Science* **63**(6): 625-644.
- Brendel O, Pot D, Plomion C, Rozenberg P, Guehl JM. 2002. Genetic parameters and QTL analysis of  $\delta^{13}\text{C}$  and ring width in maritime pine. *Plant, Cell & Environment* **25**(8): 945-953.
- Brendel O, Thiec D, Scotti-Saintagne C, Bodénès C, Kremer A, Guehl J-M. 2008. Quantitative trait loci controlling water use efficiency and related traits in *Quercus robur* L. *Tree Genetics & Genomes* **4**(2): 263-278.
- Broman KW, Wu H, Sen S, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**(7): 889-890.
- Casasoli M, Pot D, Plomion C, Monteverdi MC, Barreneche T, Lauteri M, Villani F. 2004. Identification of QTLs affecting adaptive traits in *Castanea sativa* Mill. *Plant, Cell & Environment* **27**(9): 1088-1101.



- Cernusak LA, Tcherkez G, Keitel C, Cornwell WK, Santiago LS, Knohl A, Barbour MM, Williams DG, Reich PB, Ellsworth DS, *et al.* 2009. Why are non-photosynthetic tissues generally  $^{13}\text{C}$  enriched compared with leaves in  $\text{C}_3$  plants? Review and synthesis of current hypotheses. *Functional Plant Biology* **36**(3): 199-213.
- Churchill GA, Doerge RW. 1994. Empirical Threshold Values for Quantitative Trait Mapping. *Genetics* **138**(3): 963-971.
- Comstock JP, McCouch SR, Martin BC, Tauer CG, Vision TJ, Xu Y, Pausch RC. 2005. The effects of resource availability and environmental conditions on genetic rankings for carbon isotope discrimination during growth in tomato and rice. *Functional Plant Biology* **32**(12): 1089-1105.
- Condon A, Richards R. 1992. Broad sense heritability and genotype x environment interaction for carbon isotope discrimination in field-grown wheat. *Australian Journal of Agricultural Research* **43**(5): 921-934.
- Condon A, Richards R, Rebetzke G, Farquhar G. 2004. Breeding for high water-use efficiency. *Journal of Experimental Botany* **55**(407): 2447-2460.
- Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**(18): 3674-3676.
- Costantini L, Battilana J, Lamaj F, Fanizza G, Grando M. 2008. Berry and phenology-related traits in grapevine (*Vitis vinifera* L.): From Quantitative Trait Loci to underlying genes. *BMC Plant Biology* **8**(1): 38.
- Cumbie WP, Eckert A, Wegrzyn J, Whetten R, Neale D, Goldfarb B. 2011. Association genetics of carbon isotope discrimination, height and foliar nitrogen in a natural population of *Pinus taeda* L. *Heredity* **107**(2): 105-114.
- Day ME, Greenwood MS, White AS. 2001. Age-related changes in foliar morphology and physiology in red spruce and their influence on declining photosynthetic rates and productivity with tree age. *Tree Physiology* **21**(16): 1195-1204.
- de Lorenzo L, Merchan F, Laporte P, Thompson R, Clarke J, Sousa C, Crespi M. 2009. A novel plant leucine-rich repeat receptor kinase regulates the response of *Medicago truncatula* roots to salt stress. *Plant Cell* **21**(2): 668-680.
- Dillen SY, Marron N, Koch B, Ceulemans R. 2008. Genetic Variation of Stomatal Traits and Carbon Isotope Discrimination in Two Hybrid Poplar Families (*Populus deltoides* 'S9-2'  $\times$  *P. nigra* 'Ghoy' and *P. deltoides* 'S9-2'  $\times$  *P. trichocarpa* 'V24'). *Annals of Botany* **102**(3): 399-407.
- Donovan LA, Ehleringer JR. 1992. Contrasting water-use patterns among size and life-history classes of a semi-arid shrub. *Functional Ecology* **6**(4): 482-488.
- Dye PJ. 2000. Water use efficiency in South African *Eucalyptus* plantations: A review. *The Southern African Forestry Journal* **189**(1): 17-26.
- Eldridge KG, Davidson J, Harwood CE, Wyk Gv. 1993. *Eucalypt domestication and breeding*: Clarendon Press.
- Elhiti M, Hebelstrup KH, Wang A, Li C, Cui Y, Hill RD, Stasolla C. 2013. Function of type-2 Arabidopsis hemoglobin in the auxin-mediated formation of embryogenic cells during morphogenesis. *The Plant Journal* **74**(6): 946-958.
- FAO. 2006. *Global forest resources assessment 2005 : progress towards sustainable forest management*. Rome: Food and Agriculture Organization of the United Nations.
- FAO. 2007. *State of the world's forests: 2007*. Rome: Food and Agriculture Organization of the United Nations.



- Farquhar G, Hubick K, Condon A, Richards R 1989.** Carbon Isotope Fractionation and Plant Water-Use Efficiency. In: Rundel PW, Ehleringer JR, Nagy KA eds. *Stable Isotopes in Ecological Research*: Springer New York, 21-40.
- Farquhar G, Richards R. 1984.** Isotopic Composition of Plant Carbon Correlates With Water-Use Efficiency of Wheat Genotypes. *Functional Plant Biology* **11**(6): 539-552.
- Forrester DI, Collopy JJ, Morris JD. 2010.** Transpiration along an age series of *Eucalyptus globulus* plantations in southeastern Australia. *Forest Ecology and Management* **259**(9): 1754-1760.
- Gion J-M, Carouche A, Deweer S, Bedon F, Pichavant F, Charpentier J-P, Bailleres H, Rozenberg P, Carocha V, Ognouabi N, et al. 2011.** Comprehensive genetic dissection of wood properties in a widely-grown tropical tree: *Eucalyptus*. *BMC Genomics* **12**(1): 301.
- Gonçalves JLD, Alvares CA, Higa AR, Silva LD, Alfenas AC, Stahl J, Ferraz SFdB, Lima WdP, Brancalion PHS, Hubner A, et al. 2013.** Integrating genetic and silvicultural strategies to minimize abiotic and biotic constraints in Brazilian eucalypt plantations. *Forest Ecology and Management* **301**(0): 6-27.
- Gonzalez-Martinez SC, Huber D, Ersoz E, Davis JM, Neale DB. 2008.** Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. *Heredity* **101**(1): 19-26.
- Grattapaglia D. 2007.** *Marker-assisted selection in Eucalyptus*. Rome, Italy: Food and Agriculture Organization of the United Nations (FAO).
- Grattapaglia D, Kirst M. 2008.** *Eucalyptus* applied genomics: from gene sequences to breeding tools. *New Phytologist* **179**(4): 911-929.
- Grattapaglia D, Sederoff R. 1994.** Genetic Linkage Maps of *Eucalyptus grandis* and *Eucalyptus urophylla* Using a Pseudo-Testcross: Mapping Strategy and RAPD Markers. *Genetics* **137**(4): 1121-1137.
- Haley CS, Knott SA. 1992.** A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**(4): 315-324.
- Hubbard RM, Bond BJ, Ryan MG. 1999.** Evidence that hydraulic conductance limits photosynthesis in old *Pinus ponderosa* trees. *Tree Physiology* **19**(3): 165-172.
- Hubbard RM, Stape J, Ryan MG, Almeida AC, Rojas J. 2010.** Effects of irrigation on water use and water use efficiency in two fast growing *Eucalyptus* plantations. *Forest Ecology and Management* **259**(9): 1714-1721.
- Hubick K, Shorter R, Farquhar G. 1988.** Heritability and Genotype x Environment Interactions of Carbon Isotope Discrimination and Transpiration Efficiency in Peanut (*Arachis hypogaea* L.). *Functional Plant Biology* **15**(6): 799-813.
- IPCC. 2007.** *IPCC fourth assessment report climate change 2007*. Geneva: Intergovernmental Panel on Climate Change.
- Ishimaru K, Yano M, Aoki N, Ono K, Hirose T, Lin SY, Monna L, Sasaki T, Ohsugi R. 2001.** Toward the mapping of physiological and agronomic characters on a rice function map: QTL analysis and comparison between QTLs and expressed sequence tags. *Theoretical and Applied Genetics* **102**(6-7): 793-800.
- Jansen RC, Stam P. 1994.** High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136**(4): 1447-1455.
- Jeon BW, Hwang J-U, Hwang Y, Song W-Y, Fu Y, Gu Y, Bao F, Cho D, Kwak JM, Yang Z, et al. 2008.** The Arabidopsis Small G Protein ROP2 Is Activated by Light in Guard Cells and Inhibits Light-Induced Stomatal Opening. *The Plant Cell Online* **20**(1): 75-87.





- Johnsen KH, Flanagan LB, Huber DA, Major JE. 1999.** Genetic variation in growth, carbon isotope discrimination, and foliar N concentration in *Picea mariana*: analyses from a half-diallel mating design using field-grown trees. *Canadian Journal of Forest Research* **29**(11): 1727-1735.
- Jung C, Seo JS, Han SW, Koo YJ, Kim CH, Song SI, Nahm BH, Choi YD, Cheong J-J. 2008.** Overexpression of AtMYB44 Enhances Stomatal Closure to Confer Abiotic Stress Tolerance in Transgenic Arabidopsis. *Plant Physiology* **146**(2): 623-635.
- Koch GW, Sillett SC, Jennings GM, Davis SD. 2004.** The limits to tree height. *Nature* **428**(6985): 851-854.
- Kottek M, Grieser J, rgen, Beck C, Rudolf B, Rubel F. 2006.** World Map of the Koppen-Geiger climate classification updated. *Meteorologische Zeitschrift* **15**(3): 259-263.
- Lambrides CJ, Chapman SC, Shorter R. 2004.** Genetic Variation for Carbon Isotope Discrimination in Sunflower. *Crop Sci.* **44**(5): 1642-1653.
- Lauteri M, Pliura A, Monteverdi MC, Brugnoli E, Villani F, Eriksson G. 2004.** Genetic variation in carbon isotope discrimination in six European populations of *Castanea sativa* Mill. originating from contrasting localities. *Journal of Evolutionary Biology* **17**(6): 1286-1296.
- Laza MR, Kondo M, Ideta O, Barlaan E, Imbe T. 2006.** Identification of Quantitative Trait Loci for  $\delta^{13}\text{C}$  and Productivity in Irrigated Lowland Rice. *Crop Sci.* **46**(2): 763-773.
- Le Roux D, Stock WD, Bond WJ, Maphanga D. 1996.** Dry mass allocation, water use efficiency and  $\delta^{13}\text{C}$  in clones of *Eucalyptus grandis*, *E. grandis*  $\times$  *camaldulensis* and *E. grandis*  $\times$  *nitens* grown under two irrigation regimes. *Tree Physiology* **16**(5): 497-502.
- Leavitt SW, Danzer SR. 1993.** Method for batch processing small wood samples to holocellulose for stable-carbon isotope analysis. *Analytical Chemistry* **65**(1): 87-89.
- Li-Beisson Y, Shorosh B, Beisson F, Andersson MX, Arondel V, Bates PD, Baud S, Bird D, DeBono A, Durrett TP, et al. 2013.** Acyl-Lipid Metabolism. *The Arabidopsis Book*: e0161.
- Liu C, Yin H, Gao P, Hu X, Yang J, Liu Z, Fu X, Luo D. 2013.** PHOSPHATIDYL SERINE SYNTHASE1 is Required for Inflorescence Meristem and Organ Development in Arabidopsis. *Journal of Integrative Plant Biology* **55**(8): 682-695.
- Lü S, Zhao H, Des Marais DL, Parsons EP, Wen X, Xu X, Bangarusamy DK, Wang G, Rowland O, Juenger T, et al. 2012.** Arabidopsis ECERIFERUM9 Involvement in Cuticle Formation and Maintenance of Plant Water Status. *Plant Physiology* **159**(3): 930-944.
- Manichaikul A, Dupuis J, Sen S, Broman KW. 2006.** Poor Performance of Bootstrap Confidence Intervals for the Location of a Quantitative Trait Locus. *Genetics* **174**(1): 481-489.
- Mareschal L, Nzila JDD, Turpault MP, Thongo M'Bou A, Mazoumbou JC, Bouillet JP, Ranger J, Laclau JP. 2011.** Mineralogical and physico-chemical properties of Ferralic Arenosols derived from unconsolidated Plio-Pleistocene deposits in the coastal plains of Congo. *Geoderma* **162**(1-2): 159-170.
- Margarido GRA, Souza AP, Garcia AAF. 2007.** OneMap: software for genetic mapping in outcrossing species. *Hereditas* **144**(3): 78-79.
- Marguerit E, Bouffier L, Chancerel E, Costa P, Lagane F, Guehl J, Plomion C, Brendel O. 2014.** The genetics of water-use efficiency and its relation to growth in maritime pine *Journal of Experimental Botany* in press.
- Marguerit E, Brendel O, Lebon E, Van Leeuwen C, Ollat N. 2012.** Rootstock control of scion transpiration and its acclimation to water deficit are controlled by different genes. *New Phytologist* **194**(2): 416-429.
- Masle J, Gilmore SR, Farquhar GD. 2005.** The ERECTA gene regulates plant transpiration efficiency in Arabidopsis. *Nature* **436**(7052): 866-870.



- Mollinari M, Margarido GRA, Vencovsky R, Garcia AAF. 2009.** Evaluation of algorithms used to order markers on genetic maps. *Heredity* **103**(6): 494-502.
- Monclus R, Dreyer E, Delmotte FM, Villar M, Delay D, Boudouresque E, Petit J-M, Marron N, Bréchet C, Brignolas F. 2005.** Productivity, leaf traits and carbon isotope discrimination in 29 *Populus deltoides* × *P. nigra* clones. *New Phytologist* **167**(1): 53-62.
- Monclus R, Dreyer E, Villar M, Delmotte FM, Delay D, Petit J-M, Barbaroux C, Le Thiec D, Bréchet C, Brignolas F. 2006.** Impact of drought on productivity and water use efficiency in 29 genotypes of *Populus deltoides* × *Populus nigra*. *New Phytologist* **169**(4): 765-777.
- Monclus R, Leple J-C, Bastien C, Bert P-F, Villar M, Marron N, Brignolas F, Jorge V. 2012.** Integrating genome annotation and QTL position to identify candidate genes for productivity, architecture and water-use efficiency in *Populus* spp. *BMC Plant Biology* **12**(1): 173.
- Nguyen-Queyrens A, Ferhi A, Loustau D, Guehl J-M. 1998.** Within-ring  $\delta^{13}\text{C}$  spatial variability and interannual variations in wood cellulose of two contrasting provenances of *Pinus pinaster*. *Canadian Journal of Forest Research* **28**(5): 766-773.
- Niinemets Ü. 2002.** Stomatal conductance alone does not explain the decline in foliar photosynthetic rates with increasing tree age and size in *Picea abies* and *Pinus sylvestris*. *Tree Physiology* **22**(8): 515-535.
- Nouvellon Y, Laclau J-P, Epron D, Kinana A, Mabiala A, Roupsard O, Bonnefond J-M, le Maire G, Marsden C, Bontemps J-D, et al. 2010.** Within-stand and seasonal variations of specific leaf area in a clonal *Eucalyptus* plantation in the Republic of Congo. *Forest Ecology and Management* **259**(9): 1796-1807.
- Osório J, Osório ML, Chaves MM, Pereira JS. 1998.** Water deficits are more important in delaying growth than in changing patterns of carbon allocation in *Eucalyptus globulus*. *Tree Physiology* **18**(6): 363-373.
- Pacín M, Legris M, Casal JJ. 2013.** COP1 re-accumulates in the nucleus under shade. *The Plant Journal* **75**(4): 631-641.
- Pereira JS, Pallardy S 1989.** Water Stress Limitations to Tree Productivity. In: Pereira JS, Landsberg JJ eds. *Biomass Production by Fast-Growing Trees*: Springer Netherlands, 37-56.
- Phillips OL, van der Heijden G, Lewis SL, Lopez-Gonzalez G, Aragao LE, Lloyd J, Malhi Y, Monteagudo A, Almeida S, Davila EA, et al. 2010.** Drought-mortality relationships for tropical forests. *New Phytol* **187**(3): 631-646.
- Piñeiro M, Gómez-Mena C, Schaffer R, Martínez-Zapater JM, Coupland G. 2003.** EARLY BOLTING IN SHORT DAYS Is Related to Chromatin Remodeling Factors and Regulates Flowering in *Arabidopsis* by Repressing FT. *The Plant Cell Online* **15**(7): 1552-1562.
- Pita P, Canas I, Soria F, Ruiz F, Toval G. 2005.** Use of physiological traits in tree breeding for improved yield in drought-prone environments. The case of *Eucalyptus globulus*. *Investigacion Agraria, Sistemas y Recursos Forestales* **14**(3): 383-393.
- Prasolova NV, Xu Z, Farquhar GD, Saffigna PG, Dieters MJ. 2001.** Canopy carbon and oxygen isotope composition of 9-year-old hoop pine families in relation to seedling carbon isotope composition, growth, field growth performance, and canopy nitrogen concentration. *Canadian Journal of Forest Research* **31**(4): 673-681.
- Prasolova NV, Xu ZH, Lundkvist K, Farquhar GD, Dieters MJ, Walker S, Saffigna PG. 2003.** Genetic variation in foliar carbon isotope composition in relation to tree growth and foliar nitrogen concentration in clones of the F1 hybrid between slash pine and Caribbean pine. *Forest Ecology and Management* **172**(2-3): 145-160.
- Price AH. 2006.** Believe it or not, QTLs are accurate! *TRENDS in Plant Science* **11**(5): 213-216.



- Rebetzke G, Condon A, Farquhar G, Appels R, Richards R. 2008.** Quantitative trait loci for carbon isotope discrimination are repeatable across environments and wheat mapping populations. *Theoretical and Applied Genetics* **118**(1): 123-137.
- Rebetzke GJ, Condon AG, Richards RA, Farquhar GD. 2002.** Selection for Reduced Carbon Isotope Discrimination Increases Aerial Biomass and Grain Yield of Rainfed Bread Wheat. *Crop Sci.* **42**(3): 739-745.
- Rencoret J, Gutiérrez A, del Río José C. 2007.** Lipid and lignin composition of woods from different eucalypt species. *Holzforschung*. 165.
- Rönnberg-Wästljung AC, Glynn C, Weih M. 2005.** QTL analyses of drought tolerance and growth for a *Salix dasyclados*  $\times$  *Salix viminalis* hybrid in contrasting water regimes. *TAG Theoretical and Applied Genetics* **110**(3): 537-549.
- Schulze E-D, Turner NC, Nicolle D, Schumacher J. 2006.** Leaf and wood carbon isotope ratios, specific leaf areas and wood growth of Eucalyptus species across a rainfall gradient in Australia. *Tree Physiology* **26**(4): 479-492.
- Scott DF, Prinsloo FW. 2008.** Longer-term effects of pine and eucalypt plantations on streamflow. *Water Resources Research* **44**(7): W00A08.
- Silveira L, Alonso J. 2009.** Runoff modifications due to the conversion of natural grasslands to forests in a large basin in Uruguay. *Hydrological Processes* **23**(2): 320-329.
- Specht JE, Chase K, Macrander M, Graef GL, Chung J, Markwell JP, Germann M, Orf JH, Lark KG. 2001.** Soybean response to water: A QTL analysis of drought tolerance. *Crop Science* **41**(2): 493-509.
- Stam P. 1993.** Construction of integrated genetic linkage maps by means of a new computer package: Join Map. *The Plant Journal* **3**(5): 739-744.
- Stape JL, Binkley D, Ryan MG. 2008.** Production and carbon allocation in a clonal Eucalyptus plantation with water and nutrient manipulations. *Forest Ecology and Management* **255**(3-4): 920-930.
- Stape JL, Binkley D, Ryan MG, Fonseca S, Loos RA, Takahashi EN, Silva CR, Silva SR, Hakamada RE, Ferreira JMD, et al. 2010.** The Brazil Eucalyptus Potential Productivity Project: Influence of water, nutrients and stand uniformity on wood production. *Forest Ecology and Management* **259**(9): 1684-1694.
- Teulat B, Merah O, Sirault X, Borries C, Waugh R, This D. 2002.** QTLs for grain carbon isotope discrimination in field-grown barley. *Theoretical and Applied Genetics* **106**(1): 118-126.
- Turlapati P, Kim K-W, Davin L, Lewis N. 2011.** The laccase multigene family in *Arabidopsis thaliana*: towards addressing the mystery of their gene function(s). *Planta* **233**(3): 439-470.
- Van Ooijen JW. 2011.** JoinMap 4.1, Software for the Calculation of Genetic Linkage Maps in Experimental Populations. Wageningen, Netherlands: Kyazma BV.
- Velikova V, Loreto F, Brilli F, Stefanov D, Yordanov I. 2008.** Characterization of juvenile and adult leaves of Eucalyptus globulus showing distinct heteroblastic development: photosynthesis and volatile isoprenoids. *Plant Biology* **10**(1): 55-64.
- Verhaegen D, Plomion C, Gion JM, Poitel M, Costa P, Kremer A. 1997.** Quantitative trait dissection analysis in Eucalyptus using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. *TAG Theoretical and Applied Genetics* **95**(4): 597-608.
- Vigneron P. 1992.** Création et amélioration de variétés hybrides d'Eucalyptus au Congo. *Bois et Forêts des Tropiques* **234**: 29-42.



- Voorrips RE. 2002.** MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *Journal of Heredity* **93**(1): 77-78.
- Weih M, Ronnberg-Wastljung AC, Glynn C. 2006.** Genetic basis of phenotypic correlations among growth traits in hybrid willow (*Salix dasyclados* x *S. viminalis*) grown under two water regimes. *New Phytologist* **170**(3): 467-477.
- White DA, Crombie DS, Kinal J, Battaglia M, McGrath JF, Mendham DS, Walker SN. 2009.** Managing productivity and drought risk in *Eucalyptus globulus* plantations in south-western Australia. *Forest Ecology and Management* **259**(1): 33-44.
- Whitehead D, Beadle CL. 2004.** Physiological regulation of productivity and water use in *Eucalyptus*: a review. *Forest Ecology and Management* **193**(1-2): 113-140.
- Wu J, Tan X, Wu C, Cao K, Li Y, Bao Y. 2013.** Regulation of Cytokinesis by Exocyst Subunit SEC6 and KEULE in *Arabidopsis thaliana*. *Molecular Plant* **6**(6): 1863-1876.
- Wu Y, Close TJ, Lonardi S. 2008.** On the accurate construction of consensus genetic maps. *Computational systems bioinformatics / Life Sciences Society. Computational Systems Bioinformatics Conference 7*: 285-296.
- Wu Y, Close TJ, Lonardi S. 2011.** Accurate Construction of Consensus Genetic Maps via Integer Linear Programming. *Ieee-Acm Transactions on Computational Biology and Bioinformatics* **8**(2): 381-394.
- Wullschlegel SD, Meinzer FC, Vertessy RA. 1998.** A review of whole-plant water use studies in tree. *Tree Physiology* **18**(8-9): 499-512.
- Xu S. 2003.** Theoretical Basis of the Beavis Effect. *Genetics* **165**(4): 2259-2268.
- Xu Z, Prasolova N, Lundkvist K, Beadle C, Leaman T. 2003.** Genetic variation in branchlet carbon and nitrogen isotope composition and nutrient concentration of 11-year-old hoop pine families in relation to tree growth in subtropical Australia. *Forest Ecology and Management* **186**(1-3): 359-371.
- Yoder B, Ryan M, Waring R, Schoettle A, Kaufmann M. 1994.** Evidence of reduced photosynthetic rates in old trees. *Forest Science* **40**(3): 513-527.
- Zeng ZB. 1994.** Precision mapping of quantitative trait loci. *Genetics* **136**(4): 1457-1468.
- Zhang J, Fins L, Marshall JD. 1994.** Stable carbon isotope discrimination, photosynthetic gas exchange, and growth differences among western larch families. *Tree Physiology* **14**(5): 531-539.
- Zheng X, Wu S, Zhai H, Zhou P, Song M, Su L, Xi Y, Li Z, Cai Y, Meng F, et al. 2013.** *Arabidopsis* Phytochrome B Promotes SPA1 Nuclear Accumulation to Repress Photomorphogenesis under Far-Red Light. *The Plant Cell Online* **25**(1): 115-133.
- Zotz G, Wilhelm K, Becker A. 2011.** Heteroblasty—A Review. *The Botanical Review* **77**(2): 109-151.





## **6 Architecture génétique de la dynamique de croissance : de la croissance totale à la réponse journalière**







## 6.1 Introduction

Les variations journalières du rayon du tronc sont étudiées depuis plusieurs décennies pour comprendre la réaction de l'arbre face à l'évolution, à court terme, des conditions environnementales [1, 2]. Le suivi de ces variations journalières est réalisé de façon automatique par des dendromètres, et permet ainsi d'obtenir une très grande résolution temporelle [2]. Le signal obtenu par cette méthode traduit plusieurs phénomènes : le rétrécissement puis le gonflement des tissus périphériques, liés au déstockage et stockage d'eau du tronc, la croissance du cambium ainsi que la dilatation thermique des tissus [1-10]. De ce fait, les contractions journalières du rayon du tronc ont été reliées à plusieurs variables environnementales, notamment la température de l'air, les précipitations, la teneur en eau du sol et la demande évaporatoire [5, 6, 11, 12]. Par ailleurs, l'utilisation de dendromètres pour mesurer la croissance à l'échelle de la journée reste controversée. En effet, l'augmentation journalière du rayon du tronc, mesurée par des dendromètres, peut être imprécise du fait que le rétrécissement ou le gonflement du tronc est réversible, ce qui induit un biais dans la croissance estimée. L'intégration du signal journalier à différentes échelles de temps peut néanmoins permettre une meilleure compréhension de la dynamique de croissance [13]. Bien que ce signal de micro-variations du rayon ait été étudié sous de nombreux angles pour plusieurs espèces, les bases génétiques de cette réponse "instantanée" à l'environnement n'ont pour l'instant jamais été étudiées chez les arbres.

Dans ce dernier chapitre, l'objectif est donc de mettre en lien le déterminisme génétique de la croissance avec celui de la plasticité phénotypique (sous l'angle des micro-variations du rayon), face aux contraintes environnementales au champ. Dans une première partie, les hypothèses formulées dans le chapitre 3 sont testées à partir de données de croissance ponctuelle. Le déterminisme génétique de la croissance pour les trois premières années de plantation a été étudié sur 960 descendants afin de préciser l'architecture génétique de la croissance. Dans une seconde partie, la plasticité phénotypique des paramètres du cycle journalier a été étudiée à différentes échelles de temps, et sa variabilité au sein de la descendance a été caractérisée. L'architecture génétique de ces paramètres a ensuite été analysée sous deux approches complémentaires. Ces analyses sont réalisées à différentes échelles de temps, pour 222 génotypes du croisement interspécifique de référence *E. urophylla* x *E. grandis*.

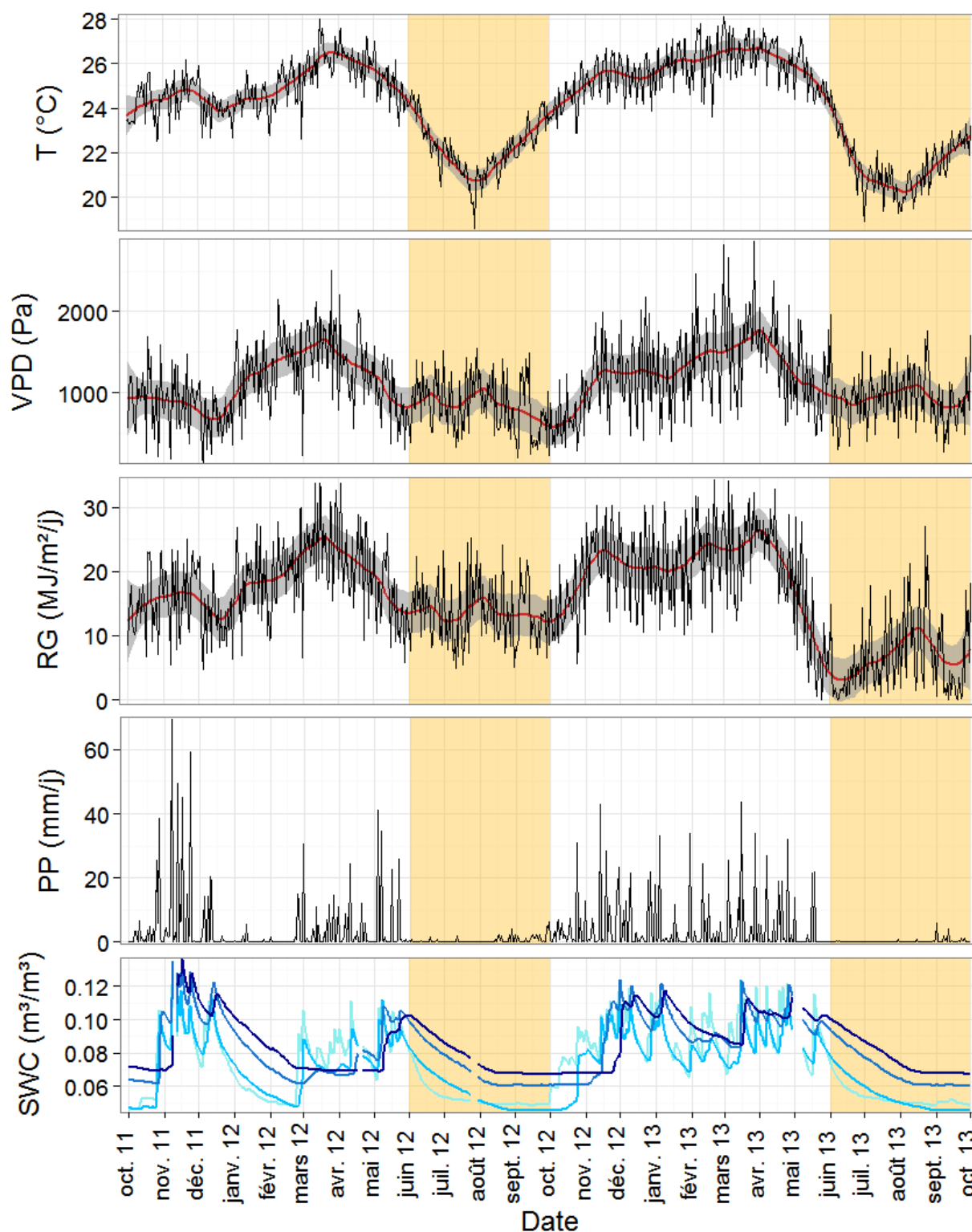


Figure 6.1 : Valeurs journalières sur les deux ans de suivi pour la température moyenne de l'air (TC), le déficit de pression de vapeur (VPD) maximum, le rayonnement global (RG), la pluviométrie (PP) et le contenu en eau du sol (SWC) de 15 cm (bleu clair) à 200 cm (bleu foncé). Pour les trois graphiques supérieurs, la courbe rouge est une moyenne mobile sur 2 semaines et la zone grisée indique l'intervalle de confiance de cette moyenne (99%). La saison sèche est représentée en orange.



## 6.2 Matériel et méthodes

Les éléments présentés dans cette partie résument des informations importantes abordées dans le chapitre 2 et détaillent les méthodes spécifiques à cette partie.

### 6.2.1 Matériel végétal et essai au champ

La descendance étudiée est issue du croisement entre *Eucalyptus urophylla* (*E.u.*, parent femelle) and *E. grandis* (*E.g.*, parent mâle). Cette descendance F1 a été installée au champ en novembre 2010 à Kissoko (4.45° S, 12° E) dans la région de Pointe-Noire en République du Congo. La densité de plantation est de 802 arbres/ha et chaque génotype est présent en un seul exemplaire. Les sols dans cette région sont principalement composés de sable ( $\approx 90\%$ ) et sont considérés comme pauvres du fait de leur faible rétention en eau et de leur faible capacité d'échange cationique [14]. Ils sont classés en tant qu'Arenosols ferraliques [15] et sont naturellement recouverts par de la savane herbacée. Le climat présente deux saisons contrastées en termes de disponibilité en eau : la saison sèche (DS) de juin à septembre, et la saison des pluies (RS) d'octobre à mai, qui concentre la majorité des précipitations (1200 mm en moyenne sur l'année).

### 6.2.2 Suivi des variables environnementales

Un suivi en continu des variables environnementales a été effectué grâce à une station météo positionnée directement sur l'essai. Ceci a permis de mesurer la température, l'humidité relative de l'air (CS215, Campbell Scientific), la pluviométrie (ARG100, Campbell Scientific) et le rayonnement global incident (CS300, Campbell Scientific). Le contenu en eau du sol (SWC en  $\text{m}^3_{\text{eau}}/\text{m}^3_{\text{sol}}$ ) a été mesuré avec des sondes réflectométriques (107, Campbell Scientific) positionnées à quatre profondeurs (15, 50, 100 and 200 cm) sur l'inter rang. La période de mesure est identique à celle du suivi des micro-variations du rayon : de début octobre 2011 à fin septembre 2013. Cette période englobe ainsi deux saisons des pluies (RS1 d'octobre 2011 à mai 2012 et RS2 d'octobre 2012 à mai 2013) et deux saisons sèches (DS1 de juin à septembre 2012 et DS2 de juin à septembre 2013) (Figure 6.1). Ces données ont permis de calculer des variables plus intégratrices pour caractériser le milieu.

Le déficit de pression de vapeur d'eau (vapor pressure deficit, VPD) a été calculé selon la formule suivante :

$$\text{VPD} = \left(1 - \frac{\text{HR}}{100}\right) \times e_s$$





$HR$  étant l'humidité relative (en %) directement mesurée sur l'essai et  $e_s$  la pression de vapeur saturante (en Pa) calculée à partir de la formule suivante :

$$e_s = (f1 + f2 \times P) \times a \times e^{(b \times T)/(c+T)}$$

Avec  $f1$  une constante (1.0007),  $f2$  une constante ( $3.46 \times 10^{-8}$ ),  $P$  la pression en Pa,  $a$  une constante (611.21),  $b$  une constante (17.502),  $c$  une constante (240.97) et  $T$  la température mesurée sur l'essai (en °C).

La réserve en eau relative du sol (REW pour relative extractable soil water) a été calculée à partir des données d'humidité du sol selon les formules suivantes :

$$REW = \sum_{i=1}^n z_i \frac{EW_i}{EW_i^M} = z_i \frac{(\theta - \theta_m)_i}{(\theta_{fc} - \theta_m)_i}$$

REW est la somme sur l'ensemble du profil, du ratio entre la réserve en eau (EW) à un instant donné et la réserve en eau maximale ou réserve utile ( $EW^M$ ), pondérée par la hauteur de la couche de sol considérée ( $z_i$ ). EW est égale à la différence entre le contenu en eau du sol à un instant donné ( $\theta$ ) et le contenu en eau du sol au point de flétrissement permanent ( $\theta_m$ ).  $EW^M$  est définie comme la différence entre le contenu en eau à la capacité au champ ( $\theta_{fc}$ ) et  $\theta_m$ . REW varie donc entre 0 (point de flétrissement permanent) et 1 (capacité au champ). Les données d'humidité du sol des quatre plateaux ont été moyennées par profondeur afin d'obtenir un profil moyen pour l'ensemble de la parcelle. Chaque profondeur de mesure est utilisée pour représenter une couche du sol. La mesure à 15 cm représente la couche de sol entre 0 et 30 cm, celle à 50 cm la couche entre 30 et 70 cm, celle à 100 cm la couche entre 70 à 130 cm, et celle à 200 cm la couche entre 130 à 270 cm. Sachant que la profondeur des racines fines peut atteindre 3 m pour une plantation de 1 an [16, 17], il a été considéré que l'ensemble du profil mesuré est prospecté par les racines pour toute la durée de l'étude. Par ailleurs, bien que dans ce type de plantation la profondeur d'enracinement soit supérieure à 3 m, le choix a été fait de ne mesurer que cette partie du profil, qui est la plus variable.

Enfin l'indice d'aridité de "de Martonne" ( $I_{DM}$ ) a été calculé selon la formule suivante :

$$I_{DM} = 12 p / (t + 10)$$

Avec  $p$ , la pluviométrie mensuelle (en mm) et  $t$ , la température moyenne mensuelle (en °C).



Hauteur	↓		↓		↓		↓	↓	↓		↓
Diamètre	↓	↓	↓	↓	↓		↓	↓	↓		↓
Suivi en continu											
Age en mois	3	6	10	15	18	22	27	30			35

Tableau 6.1 : Age de mensuration pour les mesures ponctuelles de croissance en hauteur et diamètre. La période de suivi en continu des micro-variations du rayon est également indiquée. La saison sèche est représentée en orange et la saison des pluies en bleu.

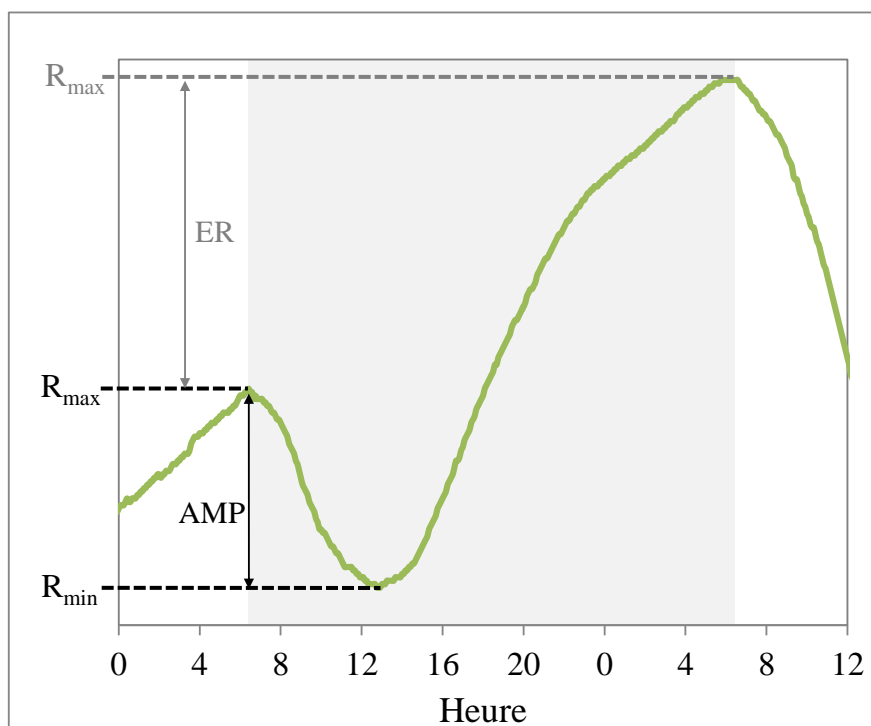


Figure 6.2 : Représentation de l'évolution du rayon au cours du cycle circadien (zone grisée). Les variables extraites de ce cycle sont le maximum ( $R_{\max}$ ) et le minimum ( $R_{\min}$ ) journalier. La différence entre les deux indique l'amplitude journalière (AMP) et la différence entre deux maximums successifs traduit l'évolution nette du rayon (ER).



### 6.2.3 Données phénotypiques : mesures ponctuelles et suivi en continu

Des mesures ponctuelles de hauteur (HT) et de diamètre à 20 cm au-dessus du sol (D) ont été effectuées à différents âges sur 960 individus (Tableau 6.1). Des accroissements entre mesures successives ont été calculés et sont notés de la manière suivante :

$$\text{Caractère } n_2 - n_1 = \text{Caractère } n_2 - \text{Caractère } n_1$$

Le *Caractère* étant la hauteur (HT) ou le diamètre (D) à l'âge  $n$ . Le coefficient de corrélation entre ces accroissements et l'indice de 'de Martonne' cumulé a été calculé pour chaque génotype. Compte tenu du faible nombre d'accroissements en hauteur disponible pour la totalité des génotypes (4), cette corrélation n'a été calculée que pour les accroissements en diamètre, en utilisant les 6 mesures disponibles. En parallèle de ces mesures ponctuelles, un suivi en continu des micro-variations du rayon a été effectué durant deux ans : de début octobre 2011 à fin septembre 2013, sur 222 arbres instrumentés. Parmi ces arbres, douze ont été supprimés de l'analyse soit parce qu'ils sont morts en court d'expérimentation, soit parce que la qualité du signal de micro-variations était insuffisante (bruit électrique). Le signal de micro-variations a été enregistré toutes les cinq minutes grâce à un micro-dendromètre positionné sur le tronc. Le micro-dendromètre est maintenu par un support composé d'une armature en aluminium et de deux tiges en acier fichées dans le tronc (cf chapitre 2). Les données de micro-variations brutes ont subi plusieurs traitements avant d'être analysées. Ces traitements sont décrits dans le paragraphe 2.3.4. Deux paramètres ont été extraits du cycle circadien (période entre deux maximums successifs) : l'amplitude journalière (AMP) et l'évolution nette du rayon (ER) (Figure 6.2). Le rythme circadien de micro-variations du rayon est connu chez les arbres pour intégrer plusieurs processus, dont deux principaux : le fonctionnement hydrique de la plante (mouvement d'eau) et la croissance du méristème secondaire (division et élongation cellulaire). Le premier serait lié à AMP et le second à ER, d'où le choix de ces deux paramètres.

Afin de tester l'effet des variables environnementales (TC, RG, VPD, PP et REW) sur la variabilité des paramètres du cycle circadien (AMP et ER) Un modèle à effet fixe a été ajusté pour chaque génotype:

$$y = \mu + TC + RG + VPD + PP + REW + VPD \times REW + \varepsilon$$







avec  $y$  le phénotype de l'arbre (AMP ou ER),  $\mu$  la moyenne générale, TC, RG, VPD, PP et REW les effets fixes des variables environnementales, VPD $\times$ REW l'interaction entre les deux variables et  $\varepsilon$  l'erreur résiduelle du modèle suivant une loi normale  $\sim N(0, \sigma_\varepsilon^2)$ . Pour ce modèle la fonction *lm* du logiciel R [18] a été utilisée.

L'ensemble des analyses sur les variables environnementales ainsi que sur les données phénotypiques a été réalisé avec le logiciel R 3.0.2 [18]. Les graphiques ont été réalisés avec le package R *ggplot2* [19].

#### 6.2.4 Cartographie génétique et détection de QTL

La stratégie du double pseudo test cross [20] a été utilisée pour la cartographie génétique des deux parents du croisement. Une carte génétique par parent a été construite à l'aide du package R *onemap* [21] et de l'algorithme RECORD [22, 23]. La construction des cartes est détaillée dans le chapitre 4. L'analyse QTL a été réalisée avec le package R *qtl* [24] en "composite interval mapping". La méthode de régression de Haley Knott [25] a été utilisée pour estimer les génotypes par intervalles de 0.5 cM, avec un taux d'erreur de génotypage de 0,01%. Pour tous les caractères, 10000 permutations ont été effectuées afin de déterminer le seuil de signification de 5% au niveau du génome [26]. L'intervalle de confiance de Bayes (BCI) à 95% a été calculé pour chaque QTL significatif. Le logiciel MapChart [27] a été utilisé pour représenter des cartes génétiques et les QTL.

### 6.3 Résultats : caractérisation de l'architecture génétique de la dynamique de croissance sur un large effectif

L'analyse QTL pour des caractères de croissance classique a été réalisée en utilisant un effectif de 960 pleins-frères. À notre connaissance, aucun résultat QTL n'a été rapporté pour ce type de traits chez les arbres avec un tel effectif. Cette partie nous a permis de préciser les variations du contrôle génétique de la croissance au jeune âge, grâce à une puissance de détection plus importante qu'avec les effectifs couramment utilisés.

#### 6.3.1 Variations et corrélations phénotypiques

Les mesures ponctuelles de croissance ont été réalisées au cours des trois premières années de croissance. Sur cette période, la croissance en hauteur est en moyenne de 17.9 m et

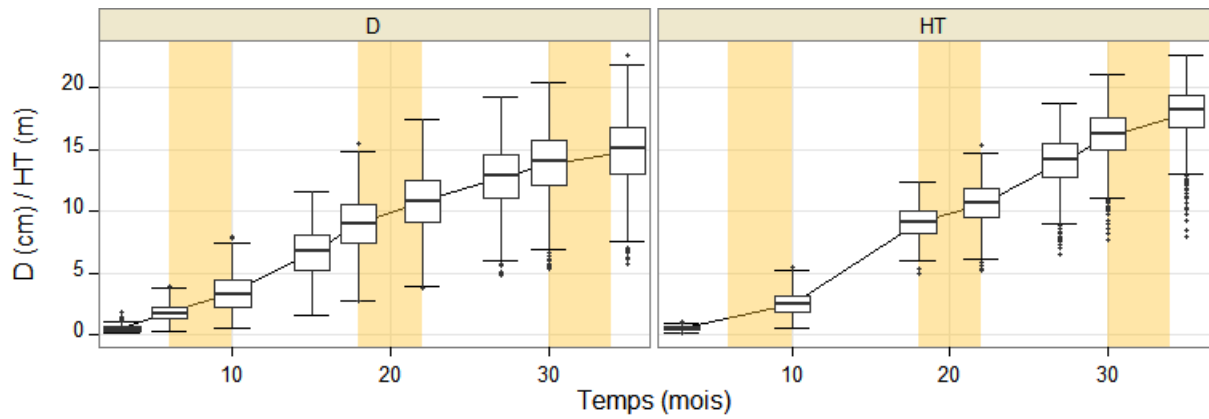


Figure 6.3 : Trajectoire de croissance radiale (D) et apicale (HT) jusqu'à trois ans. Les bandes verticales orange indiquent les saisons sèches.

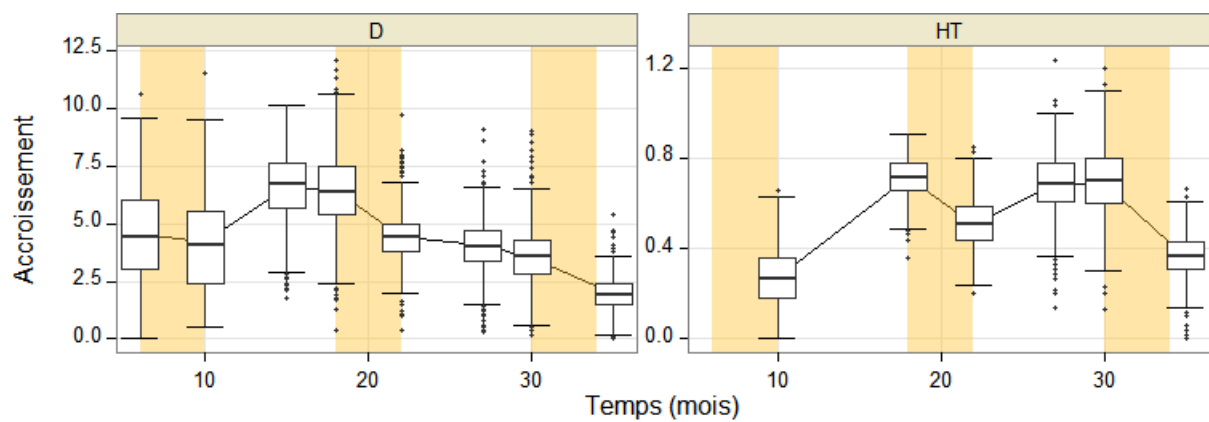


Figure 6.4 : Evolution temporelle des accroissements entre deux mesures successives de diamètre (D en cm/mois) ou de hauteur (HT en m/mois). Les bandes verticales orange indiquent les saisons sèches.

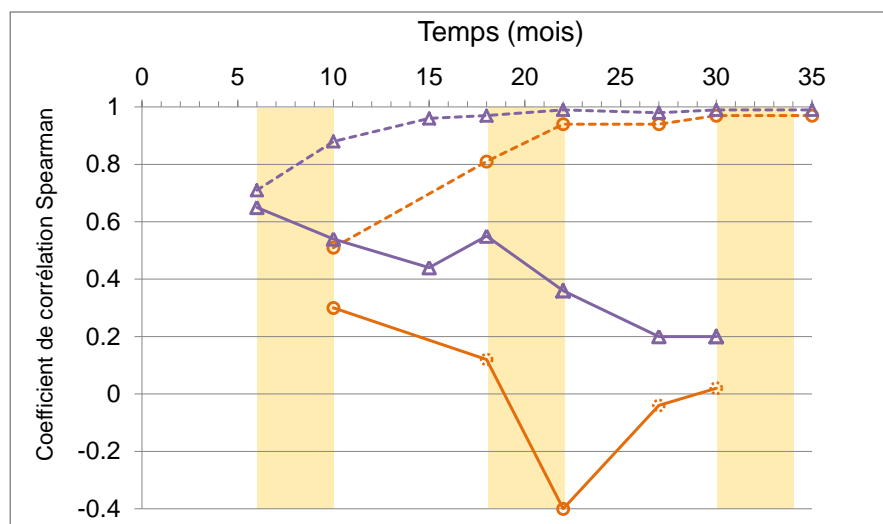


Figure 6.5 : Evolution des coefficients de corrélation « âge-âge » pour la croissance cumulée (courbes en pointillés) et pour les accroissements (courbes en traits pleins). Le diamètre est représenté en violet avec des triangles et la hauteur en orange avec des cercles. Les bandes verticales orange indiquent les saisons sèches.



celle en diamètre de 14.8 cm (Tableau S6.1, Figure 6.3). À titre de comparaison, la hauteur moyenne sur l'essai P98 planté à la même période (en novembre) et à la même densité était de 16.2 m à 36 mois. Il ne se dégage pas de tendance claire au niveau de l'évolution temporelle des accroissements pour HT (Figure 6.4). On observe cependant une réduction des accroissements pendant les saisons sèches pour HT, celle-ci étant moins marquée pour D (Figure 6.4). En effet, pour D une réduction progressive des accroissements est visible après 18 mois. Cette diminution est accentuée lors des deux saisons sèches. La variabilité de HT et celle de D tendent à diminuer progressivement au cours du temps. En effet, sur la période de mesure, le coefficient de variation phénotypique (CVp) passe de 0.37 à 0.13 pour HT et de 0.43 à 0.2 pour D. Par contre, la tendance pour les accroissements n'est pas la même : après une diminution brutale après 10 mois, le CVp augmente de manière régulière jusqu'à 35 mois à la fois pour D et HT (Tableau S6.1). La variabilité est donc marquée avant 10 mois et diminue fortement pendant la première saison sèche (de 11 à 15 mois).

Les coefficients de corrélation entre mesures successives présentent une augmentation régulière dans le temps de 0.71 à 0.99 pour D et de 0.51 à 0.97 pour HT (Tableau S6.2, Figure 6.5). Les corrélations initiales plus faibles soulignent que les performances à la mise en place (3 mois) ne déterminent qu'en partie la croissance future et notamment pour la hauteur. Ces fortes corrélations âge-âge contrastent avec celles entre accroissements successifs (Figure 6.5). Pour D les valeurs sont comprises entre 0.65 et 0.2, et pour HT entre 0.3 et -0.4. Le classement des génotypes sur la base des données de croissance cumulée ne reflète donc pas les changements de classement au niveau des accroissements. Ce résultat est cohérent avec ceux obtenus à partir des dispositifs P97 et P98 (cf. chapitre 3).

Le coefficient de corrélation entre l'indice de 'de Martonne' ( $I_{DM}$ ) et les accroissements en diamètre sur l'ensemble des génotypes est de 0.4 ( $p < 0.05$ ). Pour chaque génotype pris séparément, les coefficients de corrélation varient entre -0.08 et 0.94, la moyenne étant de 0.46 (Figure S6.2). Parmi ces corrélations, très peu sont significatives (2.1%,  $p < 0.05$ ). L'effet des changements de disponibilité en eau sur la croissance est donc moins important que celui mis en évidence sur le dispositif P97 (cf. chapitre 3). Cette différence peut s'expliquer par un plus faible nombre de mesure n'autorisant pas une caractérisation fine de la dynamique de croissance. Cette corrélation entre accroissement en diamètre et  $I_{DM}$  a été utilisée en analyse QTL.

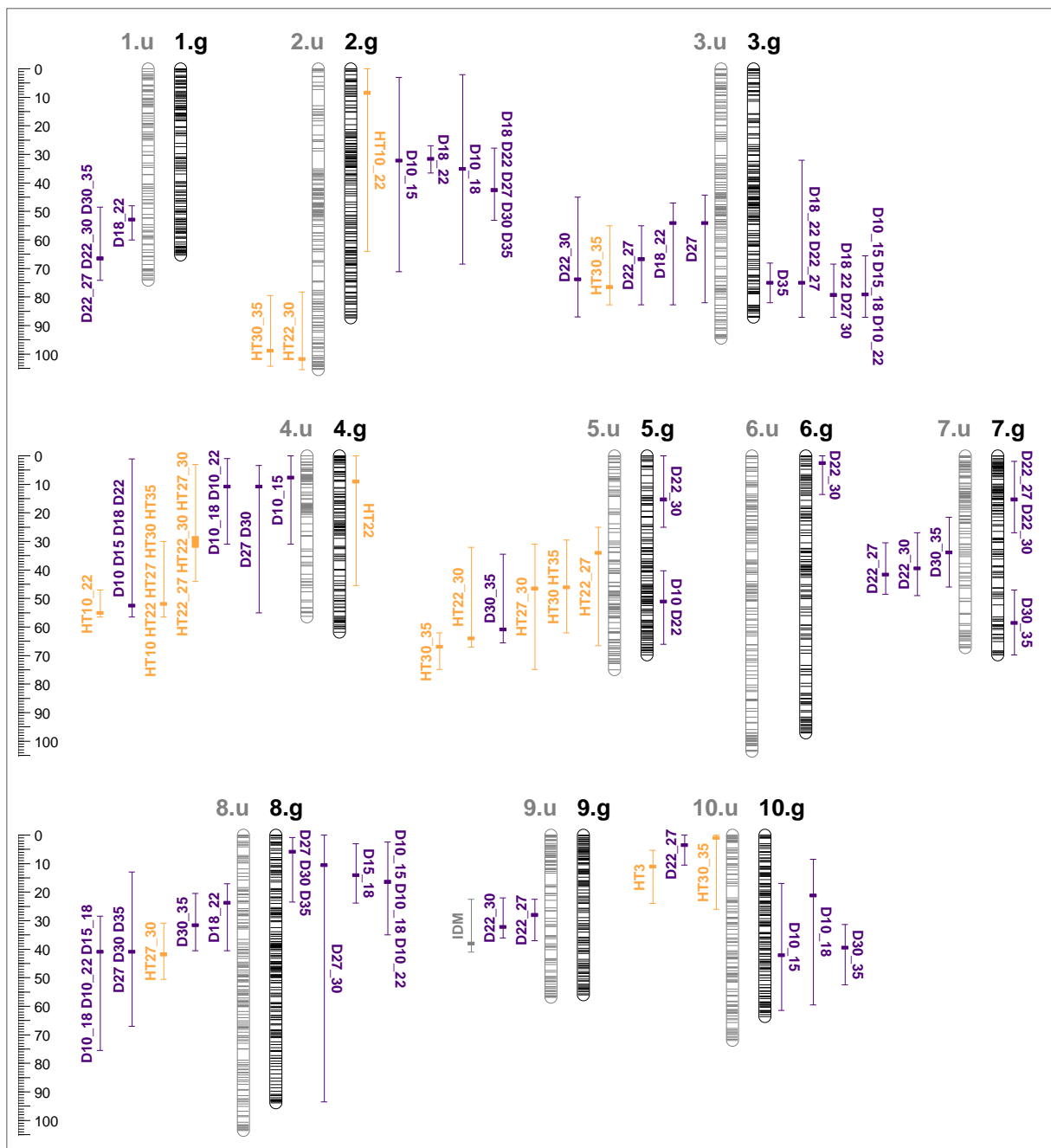


Figure 6.6 : Localisation des QTL détectés pour la croissance total pour les deux parents *E.g* (g) et *E.u* (u). HT est présentée en orange et D en violet. Lorsque plusieurs QTL colocalisent pour un même caractère mais à des âges différents, un seul QTL est positionné avec l'intervalle de confiance le plus large. Le seuil de significativité des QTL est de 5% au niveau du génome.



### 6.3.2 Architecture génétique des caractères de croissance

Un total de 55 et 41 QTL associés à 27 et 17 caractères différents ont été mis en évidence pour le parent *E.u* et le parent *E.g* respectivement (Figure 6.6, Tableau S6.3). Ces QTL sont répartis sur 9 LG (1, 2, 3, 4, 5, 7, 8, 9 et 10) pour *E.u* et 8 (LG 2, 3, 4, 5, 6, 7, 8 et 10) pour *E.g*. Seul le LG11 ne présente pas de QTL de croissance, ce qui confirme les résultats obtenus à partir de P93 et P97. La variabilité présente chez les parents du croisement ne permet pas de mettre en évidence des régions génomiques impliquées dans le déterminisme génétique de la croissance sur ce LG. Les QTL détectés expliquent un très faible pourcentage de la variance phénotypique. L'ensemble des QTL associés à un caractère donné expliquent au maximum 10.1% pour *E.u* et 6.8% pour *E.g*. En effet, un QTL pris séparément explique en moyenne 1.9% (*E.u*) et 1.7% (*E.g*) de la variance phénotypique. En comparaison, les QTL détectés sur P93 et P97 expliquent en moyenne 8% la variance phénotypique, une valeur surestimée en lien avec l'effet Beavis. De plus, le nombre de QTL détectés par caractère à partir de ce dispositif est en moyenne plus élevé (presque le double) que celui obtenu à partir des dispositifs P93 ou P97. Il est également intéressant de noter que 91% des QTL pour HT sont détectés chez *E.u*. Le contrôle génétique pour D (croissance secondaire) est quant à lui partagé entre les deux parents (45% pour *E.u*). La variabilité génétique intra-famille apportée par le parent *E.u* influe donc de manière significative sur le contrôle génétique de la croissance apicale (HT).

En relation avec les fortes corrélations âge-âge pour la croissance cumulée, des hotspots de QTL, impliquant au minimum trois mesures de croissance cumulée pour un même caractère (HT ou D), ont été mis en évidence. Pour *E.u*, 2 hotspots localisés sur le LG4 (51-52 cM) et le LG8 (40.3 cM) rassemblent respectivement 50% et 16% des QTL de croissance cumulée. Pour *E.g*, 3 hotspots localisés sur le LG2 (42 cM), le LG3 (78.8 cM) et le LG8 (5.3 cM) rassemblent 31%, 31% et 19% des QTL de croissance cumulée. De plus, quel que soit le parent considéré, tous les QTL de croissance cumulée colocalisent avec des QTL d'accroissement (excepté pour un QTL de hauteur (HT22) sur le LG4 (*E.g*)). Des QTL spécifiques des accroissements (HT ou D) sont également présents sur de nombreux LG, que ce soit pour *E.u* (LG 1, 2, 7, 9) ou *E.g* (LG 5, 6, 7, 10). Ces QTL spécifiques soulignent la diversité du contrôle génétique associée aux accroissements, ce qui se traduit par une diversité des réponses au niveau de l'arbre. Ces réponses peuvent notamment être reliées aux changements de disponibilité en eau et à la capacité des génotypes de profiter non de la



Tableau 6.2 : Caractéristiques du climat sur la période étudiée.

	Période	RS1	DS1	RS2	DS2
Température (°C)	24.1	24.9	22.2	25.7	21.4
Humidité relative (%)	93.2	93.2	93.5	93.3	92.5
VPD (Pa)	1092	1129	856	1260	925.8
Rayonnement (MJ/m/j)	16	18	13.6	19.9	6.5
Précipitation (mm)	1878.7	846	61.6	945.3	25.6
SWC (m <sup>3</sup> /m <sup>3</sup> )					
15	0.073	0.077	0.053	0.09	0.053
50	0.07	0.074	0.055	0.082	0.054
100	0.082	0.084	0.07	0.092	0.069
200	0.085	0.086	0.078	0.093	0.078
REW (%)	47.5	50	24.8	68.4	23.5

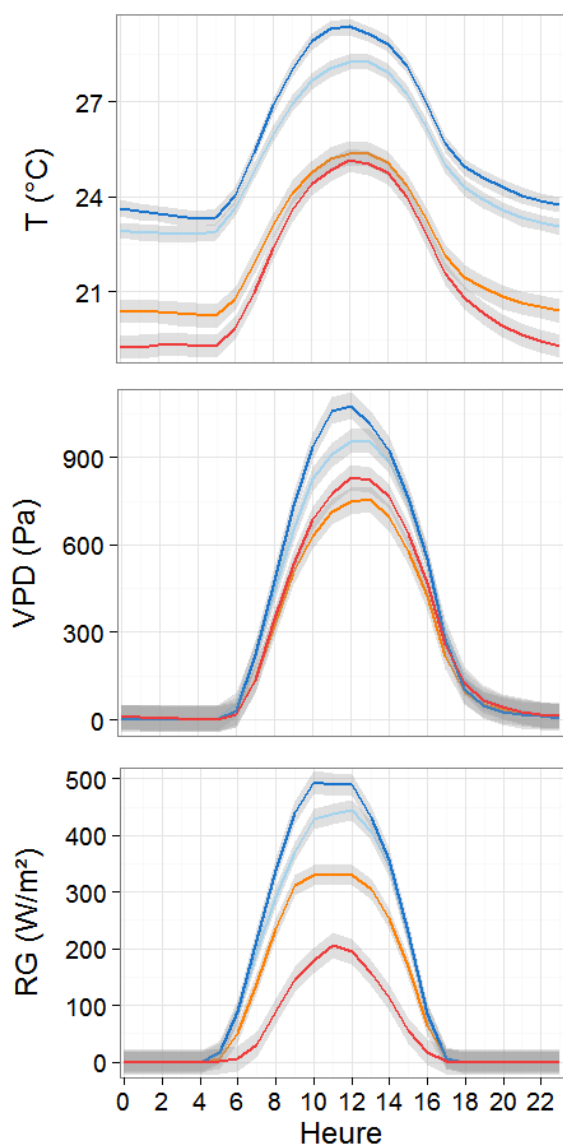


Figure 6.7 : Evolutions journalières moyennes de la température (T), du déficit de pression de vapeur (VPD) et du rayonnement global (RG). Les deux saisons des pluies (RS) sont en bleu (RS1 en clair et RS2 en foncé) et les deux saisons sèches (DS) sont en orange (DS1 en clair et DS2 en foncé). La zone grisée indique l'intervalle de confiance de la moyenne (99%).





ressource, avec comme exemple le QTL pour les coefficients de corrélation entre les accroissements en diamètre et l' $I_{DM}$  sur le LG9 (*E.u*). Dans les trois cas illustrés à la Figure S6.3, les QTL mis en évidence pour la croissance cumulée à 3 ans colocalisent avec des QTL d'accroissement à des âges antérieurs. Ces colocalisations montrent que l'architecture génétique de la croissance, à la moitié de la rotation, est en partie déterminée par la réponse à l'environnement au jeune âge.

Bien que dans cette étude, les mesures ne couvrent pas la totalité de la durée de rotation, les résultats de cette partie confirment ceux du chapitre 3. En effet, un déterminisme génétique commun entre les accroissements et la croissance cumulée a été mis en évidence, confirmant l'effet des réponses à l'environnement au jeune âge (Figure S6.3). Par ailleurs, le passage à un effectif important (960 individus) a permis de préciser l'architecture génétique de la croissance au niveau du nombre de QTL et de leurs effets. Toutefois, les intervalles de confiance restent toujours très larges (37 cM en moyenne), en lien avec l'héritabilité faible des caractères de croissance et également le manque de précision de ceux-ci (une copie par génotype). Cette augmentation d'effectif a également permis de mettre en évidence des QTL de croissance (D27, D30 et D10\_15, D10\_18) sur le LG4 d'*E.u*. Cette position correspond à celle d'un QTL de  $\delta^{13}C$  détaillée dans le chapitre 5. Or, dans ce chapitre 5, aucun QTL de croissance n'avait été détecté sur le LG4 de *E.u* avec un effectif plus faible (220 génotypes). Ce résultat est intéressant car il permet de préciser le lien entre le contrôle génétique du  $\delta^{13}C$  et celui de la croissance.

## 6.4 Résultats : La plasticité des micro-variations radiales

### 6.4.1 Caractérisation du climat

Sur les deux années de mesure, la température moyenne est de 24.1°C avec un VPD maximum moyen de 1092 Pa, et un cumul des précipitations de 1879 mm (Tableau 6.2). Sur cette période, la REW moyenne est 47.5%. Le sol est en moyenne plus sec en surface (jusqu'à 50 cm) qu'en profondeur (au-delà de 1m) avec un écart de 1% d'humidité entre ces deux couches. Le contraste entre saison des pluies (RS) et saison sèche (DS) est marqué au niveau de la disponibilité en eau avec une différence de pluviométrie de 850 mm, et une différence d'humidité de sol moyenne de 1.15% (200 cm) à 3.05% (15 cm). Ceci représente sur l'ensemble du profil une différence de REW de 35% entre les deux saisons. Ce contraste est aussi présent au niveau du rayonnement global, avec en moyenne 8.9 MJ/m<sup>2</sup> de plus par jour

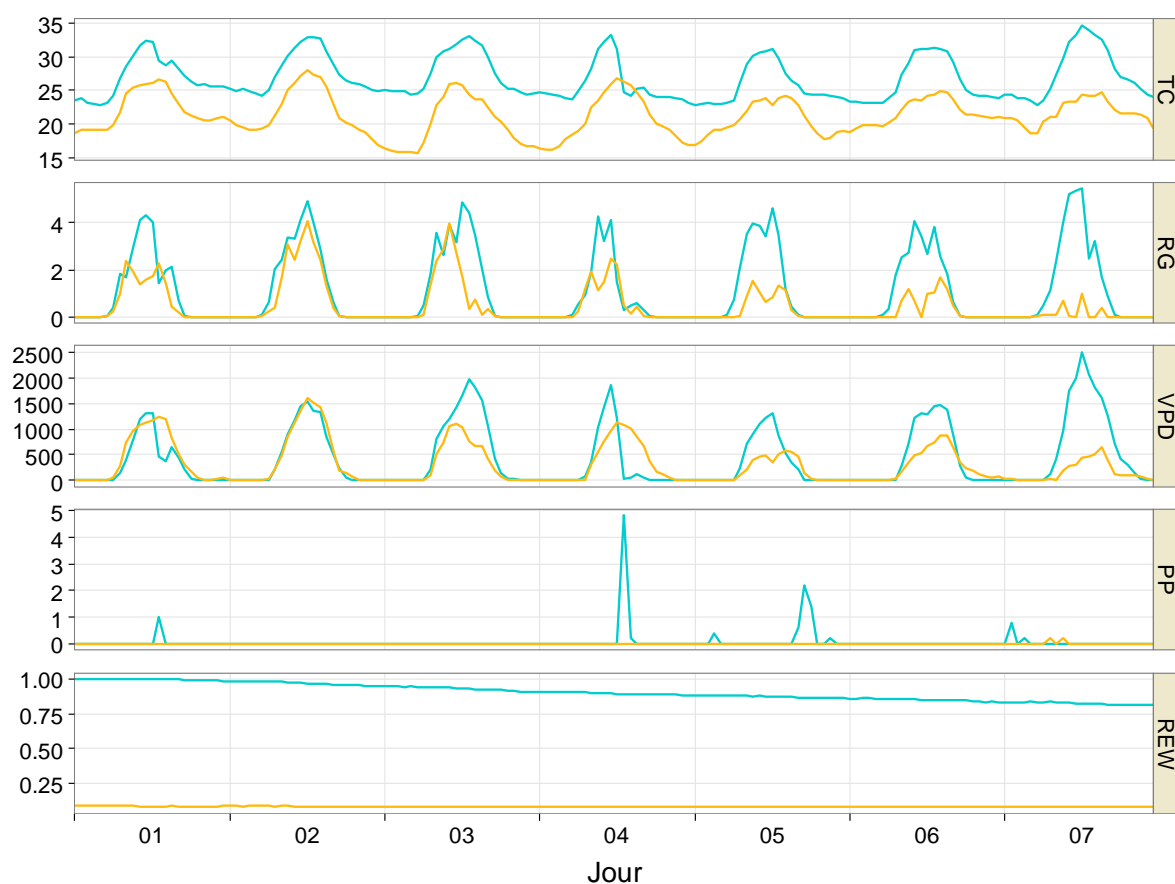


Figure 6.8 : Comparaison des variables environnementales entre deux semaines contrastées. En bleu une semaine de saison des pluies et en orange une semaine de saison sèche. Les variables représentées sont : la température moyenne de l'air ( $T$ , °C), le déficit de pression de vapeur (VPD, Pa) maximum, le rayonnement global (RG,  $\text{MJ}/\text{m}^2/\text{h}^2$ ), la pluviométrie (PP, mm) et le contenu en eau du sol (REW, %)

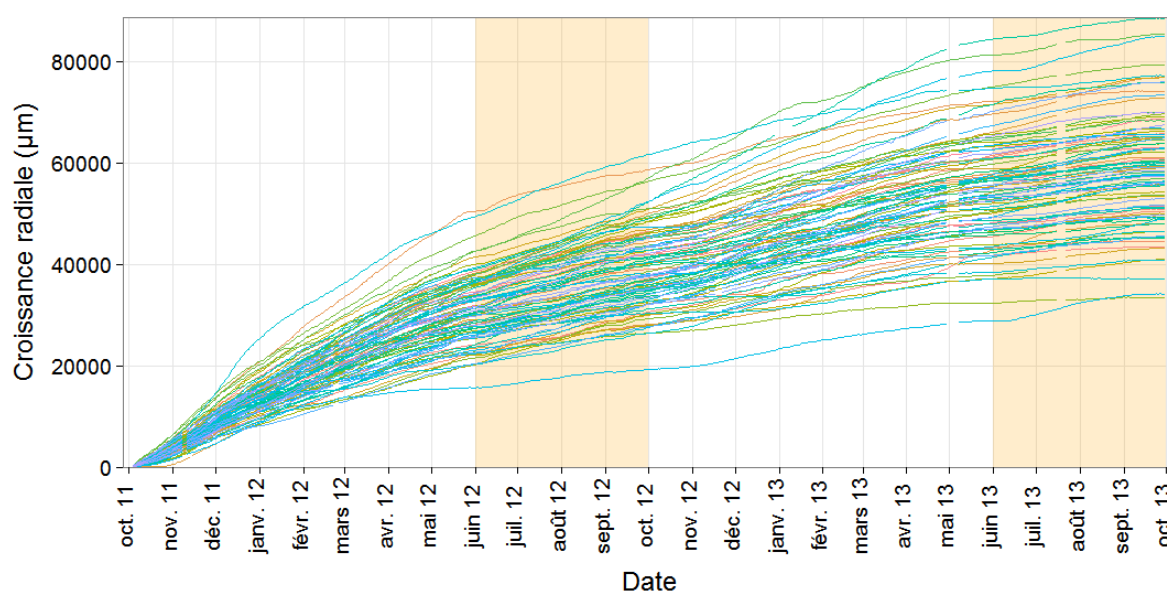


Figure 6.9 : Trajectoire de croissance à partir des mesures de micro-dendrométrie sur échantillon de 100 individus. Les bandes verticales orange indiquent les saisons sèches.



en RS qu'en DS (Figure 6.7). La différence entre RS et DS est moins marquée au niveau de la température (augmentation moyenne de 3.5°C en RS), et du VPD (déficit journalier moyen) plus important en RS (1200 Pa), qu'en DS (890 Pa). Par ailleurs, on observe une variabilité interannuelle pour un même type de saison notamment pour la pluviométrie que ce soit pour le volume (846 mm en RS1 contre 945.3 mm en RS2 et 61.6 mm en DS1 contre 25.6 mm en DS2) ou pour la répartition (petite saison sèche de janvier à février en RS1 qui n'a pas lieu en RS2), ce qui se répercute sur l'humidité du sol. En effet, la REW est plus faible en RS1 (50%) qu'en RS2 (68.4%) mais reste comparable entre DS1 (24.8%) et DS2 (23.5%).

Deux semaines, l'une en RS2 (S1) et l'autre en DS2 (S2) ont été sélectionnées sur la base des données environnementales. L'objectif est de comparer l'architecture génétique des paramètres du cycle circadien dans des conditions environnementales très contrastées. Ces deux semaines présentent donc des différences marquées pour la plupart des variables mesurées (Figure 6.8). En effet, la température moyenne est plus élevée en S1 (26.8°C) qu'en S2 (21.3 °C). Le rayonnement global est également plus élevé en S1 (27.3 Mj/m<sup>2</sup>/j) qu'en S2 (11.9 Mj/m<sup>2</sup>/j). En relation avec les deux variables précédentes, le VPD moyen est de 790 Pa en S1 et de 466 Pa en S2. Les précipitations sur ces deux semaines sont relativement faibles : 11.9 mm en S1 et 0.4 en S2. Enfin le contenu relatif en eau du sol (REW), est bien plus faible en S2 (8%) qu'en S1 (90%). La semaine S1 est caractérisée par une forte demande évaporatoire et une disponibilité en eau importante tandis que la semaine S2 présente une très faible disponibilité et une demande évaporatoire moyenne.

#### 6.4.2 Caractérisation phénotypique des paramètres du cycle circadien

La caractérisation en continu du climat nous a permis d'étudier l'impact de l'environnement sur les micro-variations du rayon. Ces effets ont été étudiés à travers deux paramètres du cycle journalier : l'amplitude journalière (AMP) et l'évolution nette du rayon (ER). Dans un premier temps, la variabilité temporelle de ces paramètres a été caractérisée dans la descendance de plein-frères. Dans un second temps, l'effet des variables environnementales sur ces paramètres est détaillé ainsi que la variabilité entre individus.

##### 6.4.2.1 Variabilité de l'amplitude journalière et de l'évolution nette du rayon

En milieu tropical, la croissance des eucalyptus est continue comme le montrent les trajectoires de croissance des arbres suivis avec un micro-dendromètre (Figure 6.9). L'accroissement moyen du rayon sur les deux années de suivi est de 56.5 mm ( $\pm$  11.3 mm)



Tableau 6.3 : Paramètres du cycle circadien (AMP et ER) moyennés sur l'ensemble des arbres pour les deux années de mesure (Période). Des moyennes (m) et des cumuls (c) par semaine et par mois sont présentés. L'écart-type associé à la moyenne est indiqué entre parenthèses.

		Période	RS1	DS1	RS2	DS2
AMP ( $\mu\text{m}$ )	Jour	95 (57.2)	88 (54.4)	89 (56.7)	101 (58.3)	102 (58.3)
	Semaine	m	93 (41.3)	87 (39.7)	85 (37.5)	99 (42.2)
		c	590 (315.9)	540 (295.5)	545 (296)	640 (317.6)
	Mois	m	92 (35.8)	86 (34.8)	86 (28.4)	99 (36.8)
		c	2558 (1162.2)	2322 (1060.2)	2398 (992.5)	2779 (1183)
					2779 (1183)	2749 (1343.8)
ER ( $\mu\text{m}$ )	Jour	78 (54.8)	122 (54.2)	66 (38.3)	64 (43.1)	33 (27)
	Semaine	m	78 (48.6)	122 (45.4)	64 (30.5)	65 (33.2)
		c	520 (338.4)	798 (350.4)	444 (216.2)	443 (236.3)
	Mois	m	78 (45.8)	121 (42.3)	66 (25.5)	64 (29.5)
		c	2252 (1344)	3417 (1330)	1944 (775.1)	1896 (894.2)
					1896 (894.2)	944 (572.1)

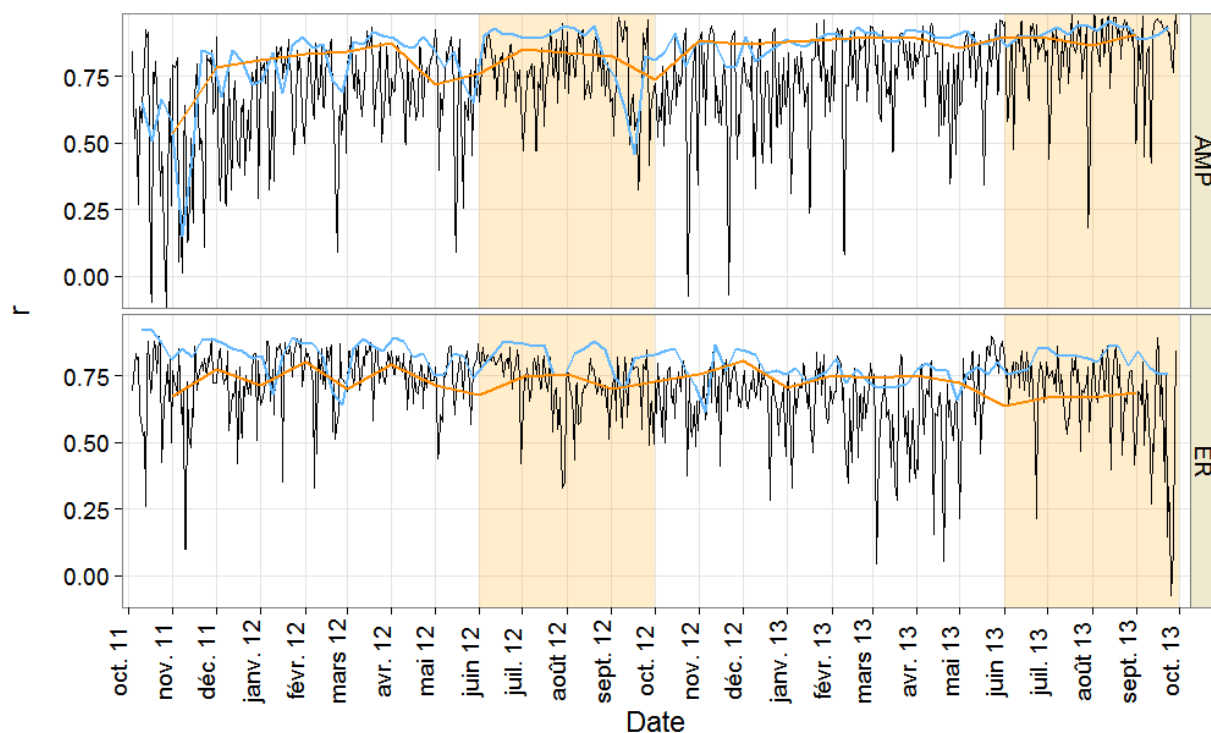


Figure 6.10 : Evolution des coefficients de corrélation entre mesures successives pour les deux paramètres du cycle circadien (AMP et ER). Les données journalières (trait noir) et les moyennes par semaine (trait bleu) et par mois (trait orange) sont représentées. Les bandes verticales orange indiquent les saisons sèches.



avec une variabilité proche de celle que l'on observe sur l'ensemble de l'essai (coefficient variation, CV de 0.2). Les deux paramètres utilisés pour décrire le cycle circadien ont une valeur moyenne sur l'ensemble de la période d'étude de 95  $\mu\text{m}$  ( $\pm 57.2$ ) pour AMP et de 78  $\mu\text{m}$  ( $\pm 54.8$ ) pour ER (Tableau 6.3). Ces valeurs moyennes sont associées à une variabilité interindividuelle et temporelle importante, avec des CV voisins de 0.5 (Figure S6.4). Les deux paramètres présentent toutefois une évolution différente au cours du temps : AMP est plutôt stable tandis qu'ER diminue. En effet, les valeurs moyennes pour AMP sont stables dans le temps : 88, 89, 101, 102  $\mu\text{m}$  pour RS1, DS1, RS2, DS2 respectivement. La variabilité associée à ces valeurs moyennes est elle aussi stable (CV  $\approx 0.6$ ). Pour ER, l'effet de la saison est plus visible avec une réduction entre RS1 (122  $\mu\text{m}$ ) et DS1 (66  $\mu\text{m}$ ) et entre RS2 (64  $\mu\text{m}$ ) et DS2 (33  $\mu\text{m}$ ) (Tableau 6.3, Figure S6.4). Ces différences s'inscrivent dans une diminution globale d'ER au cours du temps, avec une différence de 45  $\mu\text{m}$  entre RS1 et DS2 (Figure S6.4). Cette diminution s'accompagne d'une augmentation du CV de 0.1 sur les données journalières. Une tendance identique est observée pour les valeurs moyennes ou cumulées d'ER sur la semaine ou le mois.

Pour ces deux paramètres, les coefficients de corrélation entre valeurs journalières successives sont en moyenne élevés : 0.72 en moyenne pour AMP et 0.69 pour ER. Ces valeurs augmentent pour les moyennes sur la semaine et le mois : 0.84 et 0.83 pour AMP et 0.8 et 0.72 pour l'ER (Figure 6.10). Une variabilité importante est néanmoins présente pour les valeurs journalières avec des coefficients de corrélations compris entre -0.12 et 0.98 pour AMP, et entre -0.08 et 0.9 pour ER. Le classement des génotypes sur la base des données journalière est donc très variable. Ces résultats sont similaires à ceux obtenus à partir des accroissements sur les données ponctuelles de croissance, où le niveau des corrélations âge-âge est aussi très variable. Par contre, pour les paramètres moyennés sur la semaine ou le mois, les coefficients sont plus stables dans le temps. Par ailleurs, ces deux paramètres sont faiblement corrélés entre eux. En effet, pour l'ensemble de la période, les coefficients de corrélation sur l'ensemble des individus sont en moyenne de -0.12, -0.02 et 0.2 à l'échelle de la journée, de la semaine et du mois respectivement (Figure S6.5). Le classement des génotypes pour ces deux caractères est donc différent, quelle que soit l'échelle de temps considérée. Pour un même individu, les coefficients de corrélation entre ceux deux paramètres varient sur l'ensemble de la période entre -0.56 et 0.3 pour une moyenne égale à -0.22. Ces deux paramètres du cycle circadien caractérisent donc deux comportements distincts et apportent des informations complémentaires sur les génotypes.

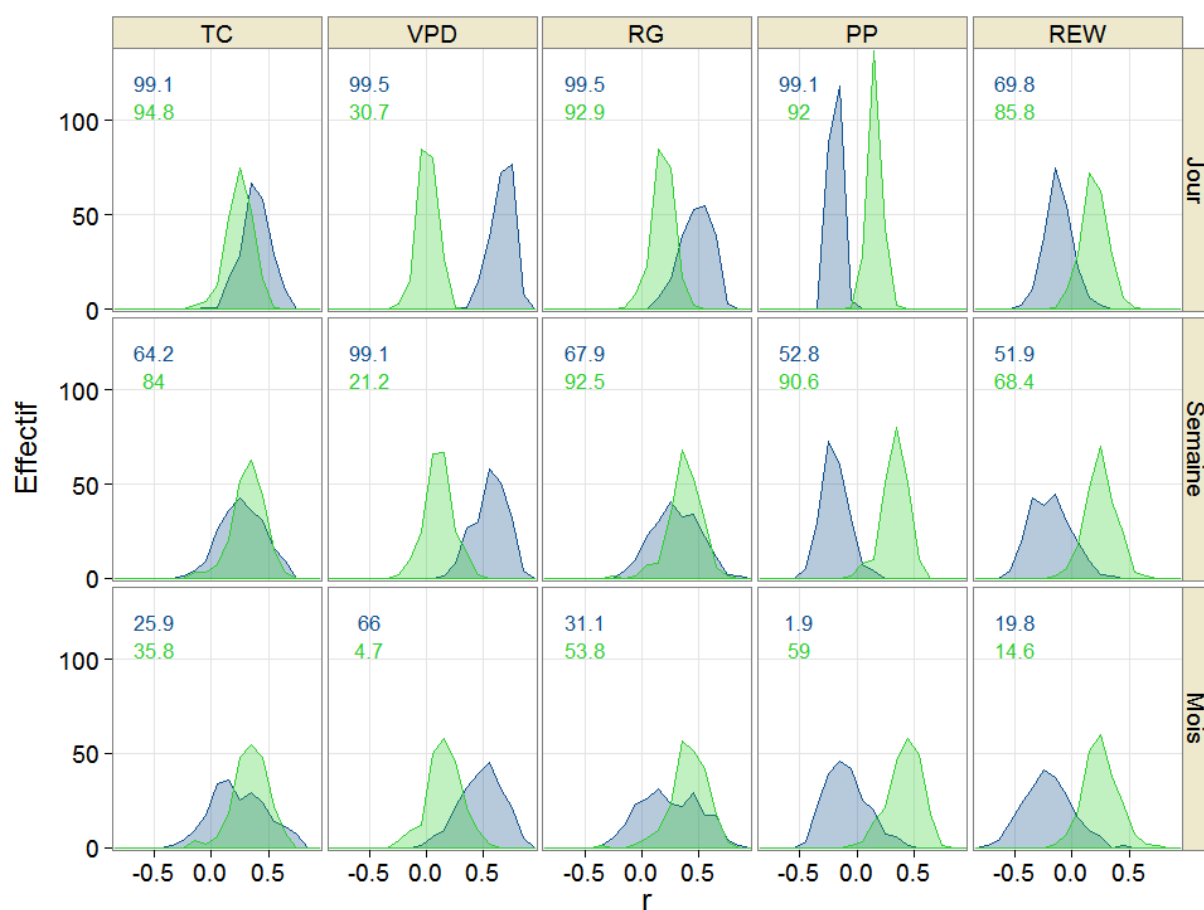


Figure 6.11 : Distribution des coefficients de corrélation de Pearson ( $r$ ) entre les paramètres (AMP en bleu ou ER en vert) et les variables environnementales sur l'ensemble de la période. Les données journalières sont utilisées pour le niveau 'Jour' et sont moyennées à la semaine ou au mois pour les niveaux 'Semaine' et 'Mois'. Les corrélations sont calculées pour chaque génotype sur toute la durée de l'expérimentation. La valeur journalière est le maximum pour TC, la moyenne pour VPD, RG et REW, et la somme pour PP. Les chiffres indiquent le pourcentage de génotypes présentant une corrélation significative.



#### 6.4.2.2 Effet des variables environnementales sur AMP et ER

Afin de caractériser la plasticité des hybrides *E. urophylla* x *E. grandis* liée aux deux paramètres du signal journalier étudiés (AMP et ER), deux approches complémentaires ont été utilisées. La première consiste à analyser les corrélations entre ces deux paramètres et les variables environnementales prises séparément. La seconde se base sur la modélisation du phénotype d'un individu, en intégrant l'ensemble des variables disponibles. Dans les deux cas, différentes échelles temps ont été analysées : les données journalières ainsi que des moyennes ou des cumuls sur la semaine et le mois.

##### Analyse des corrélations entre variables environnementales et paramètres

L'analyse des corrélations entre les variables environnementales et les paramètres journaliers, sur l'ensemble des génotypes, montre une variabilité au cours du temps pour les deux paramètres (Figure S6.6). Néanmoins, AMP présente, en moyenne pour l'ensemble des variables environnementales, des coefficients de corrélation plus forts qu'ER. Pour AMP, ils sont supérieurs en moyenne à 0.5 pour TC, VPD et RG, négatifs pour PP (-0.25) et nuls pour REW (Tableau S6.4). AMP est donc fortement corrélée à la demande atmosphérique en eau (VPD, RG) et dans une moindre mesure avec la disponibilité en eau (PP). À l'inverse, ER n'est en moyenne pas corrélée aux différentes variables environnementales avec des coefficients de corrélation proches de 0. AMP présente donc, en moyenne sur la période et sur l'ensemble des génotypes, une plasticité à l'échelle journalière plus importante qu'ER. Tous les génotypes n'ont cependant pas le même comportement face aux variations environnementales. En effet l'analyse des corrélations entre les paramètres du cycle circadien et les variables environnementales pour chaque génotype, montre une variabilité entre génotypes de ces coefficients de corrélation (Figure 6.11, Tableau S6.5). Par exemple, les coefficients de corrélation entre valeurs journalières de VPD et d'AMP varient entre 0.4 et 0.86. Cette variabilité s'accroît quand le niveau d'intégration augmente (Figure 6.11, Figure S6.7). En effet, pour les corrélations entre VPD et AMP, les coefficients de corrélation varient entre 0.14 et 0.85 pour les moyennes par semaine, et entre -0.06 et 0.88 pour les moyennes sur le mois. Des résultats similaires sont mis évidence pour ER. Le niveau d'intégration joue également sur le niveau moyen des corrélations. Pour AMP, le coefficient de corrélation moyen diminue avec l'augmentation du niveau d'intégration pour TC, VPD, RG et il reste stable pour PP et REW. À l'inverse, pour ER le coefficient de corrélation moyen augmente lorsque le niveau d'intégration augmente pour RG, PP et dans une moindre mesure pour TC et REW (Tableau S6.5). Ces deux paramètres du cycle circadien ne répondent donc pas de la





Tableau 6.4 : Pourcentage de génotypes présentant un effet significatif ( $p < 0.05$ ) des différentes variables environnementales dans la modélisation du phénotype (AMP ou ER). Les données journalières sont utilisées pour le niveau 'Jour' et sont moyennées (m) ou cumulées (c) à la semaine ou au mois pour les niveaux 'Semaine' et 'Mois'. L'interaction entre VPD et REW est également indiquée.

	AMP					ER				
	Jour	Semaine m	Semaine c	Mois m	Mois c	Jour	Semaine m	Semaine c	Mois m	Mois c
TC	33.6	18.5	40.3	5.2	17.1	35.1	14.7	19.4	1.9	9
VPD	<b>99.1</b>	<b>69.7</b>	<b>52.1</b>	25.1	16.1	40.8	<b>55.5</b>	<b>54</b>	<b>61.1</b>	<b>51.7</b>
RG	<b>74.9</b>	50.2	20.9	35.1	4.7	<b>95.7</b>	<b>84.4</b>	<b>69.2</b>	41.7	37.4
PP	20.4	21.8	13.7	4.7	0.5	37.9	16.1	9	9.5	4.7
REW	<b>65.4</b>	<b>68.2</b>	<b>63.5</b>	<b>61.1</b>	44.1	<b>98.1</b>	<b>97.2</b>	<b>94.3</b>	<b>90.5</b>	<b>87.7</b>
VPD:REW	<b>59.2</b>	<b>64.9</b>	<b>64.9</b>	<b>56.4</b>	52.1	<b>94.8</b>	<b>95.7</b>	<b>89.6</b>	<b>90.5</b>	<b>85.8</b>

m : moyenne, c : cumul

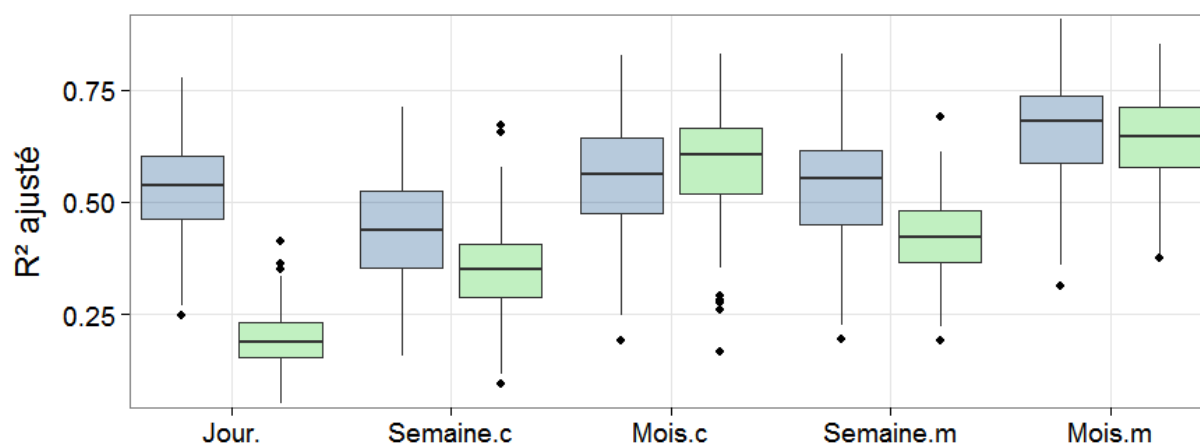


Figure 6.12 : Boxplot pour les coefficients de détermination ( $R^2$ ) des modèles par génotype en fonction du niveau d'intégration (jour, semaine, mois) pour les paramètres AMP et ER.

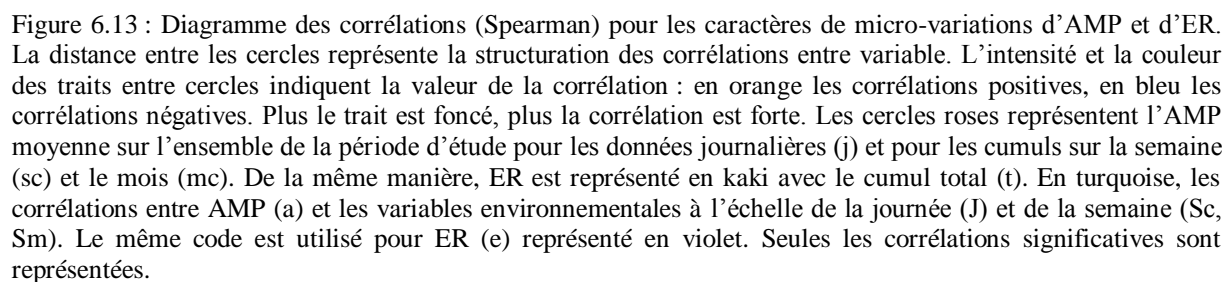


même manière aux variations environnementales, et notamment à la demande atmosphérique en eau (VPD, RG) et la disponibilité en eau (PP et REW). Cette différence de réponse peut expliquer le manque de corrélation observé entre ces deux paramètres. Les coefficients de corrélation entre les paramètres et les variables environnementales sont utilisés par la suite (analyse QTL) comme des indicateurs de la plasticité du génotype aux variations environnementales. Les corrélations à l'échelle du mois ne seront pas utilisées du fait du faible nombre de corrélations significatives.

### Modélisation du phénotype

Afin d'estimer la variabilité de l'effet des variables environnementales sur les paramètres du cycle circadien, un modèle à effet fixe pour chaque génotype, comprenant l'ensemble des variables environnementales a été ajusté. Les résultats sur le paramètre AMP montrent que pour un fort pourcentage de génotype (99.1%), l'effet du VPD est significatif à l'échelle de la journée. Lorsque l'échelle d'intégration augmente, ce pourcentage diminue, respectivement à 69.7% et à 25.1% pour les moyennes par semaine et par mois. Une tendance similaire s'observe pour les variables TC, RG et PP (Tableau 6.4). À l'inverse, le pourcentage de génotypes présentant un effet significatif de REW reste relativement stable entre les différents niveaux d'intégration avec 65.4% pour l'échelle journalière, 68.2% pour la semaine et 61.1% pour le mois. De la même manière, le pourcentage de génotype présentant un effet significatif de l'interaction entre VPD et REW est stable dans le temps, avec plus de la moitié des génotypes sur l'ensemble des niveaux d'intégration. Pour ER, l'effet de REW est significatif pour la plupart des génotypes, quelle que soit l'échelle d'intégration (87.7% pour le mois et 98.1% pour la semaine). De la même manière, l'effet de l'interaction entre VPD et REW est significatif pour la plupart des génotypes (entre 85.6% % pour le mois et 95.7% % pour la semaine). Par contre l'effet de TC et PP n'est significatif que pour un très faible pourcentage des génotypes, suggérant que leur variabilité n'explique peu ou pas la variation d'ER.

La qualité d'ajustement des modèles ( $R^2$ ) pour AMP varie entre 0.16 et 0.91 en fonction du génotype et du niveau d'intégration (Figure 6.12). Le niveau d'intégration n'a pas d'effet marqué en moyenne sur le  $R^2$ , il varie entre 0,53 (jour et semaine) et 0,66 (mois). Pour ER, sur l'ensemble des génotypes et l'ensemble des niveaux d'intégration, le  $R^2$  varie entre 0.05 et 0.85. Le  $R^2$  moyen augmente avec le niveau d'intégration, passant de 0,2 (jour) à 0,64 (mois). Le  $R^2$  est en moyenne plus faible pour les cumuls que pour les moyennes (semaine ou mois) quel que soit le paramètre. Il faut cependant rappeler que le  $R^2$  est très dépendant du modèle.





Dans notre cas, un unique modèle est utilisé pour l'ensemble des génotypes mais ne reflète pas forcément le meilleur ajustement pour chacun des génotypes.

Ces résultats de modélisation par génotype confirment donc les tendances observées avec les corrélations. Le paramètre AMP est en moyenne, pour l'ensemble des génotypes plus sensible aux effets des variables environnementales qu'ER. Cette différence, très marquée à l'échelle journalière, diminue à l'échelle de la semaine ou du mois. La modélisation permet également de souligner l'effet significatif de l'interaction entre VPD et REW sur les deux paramètres issus du cycle circadien, pour l'ensemble des échelles d'intégration.

#### 6.4.2.3 *Discussion sur la plasticité aux variations environnementales*

Les résultats de cette partie soulignent la complémentarité d'AMP et ER pour caractériser la plasticité des génotypes aux différentes variables environnementales, à différentes échelles de temps. En effet, ces deux paramètres, associés à deux composantes du signal de micro-variations du rayon (le fonctionnement hydrique pour AMP et la croissance pour ER), se sont révélés très peu corrélés entre eux. Ces fonctionnements différents sont confirmés au niveau des liens avec les variations environnementales. De manière générale, AMP est corrélée positivement au VPD et au RG, et négativement à REW, avec une interaction entre VPD et REW significative pour la moitié des génotypes. Ceci est cohérent avec ce que traduit l'AMP c.-à-d. principalement le déstockage/stockage d'eau au niveau du tronc. Ces corrélations relativement fortes à l'échelle journalière diminuent lorsque le niveau d'intégration augmente. À l'inverse, ER est peu corrélé aux variables environnementales, que ce soit à l'échelle journalière, hebdomadaire ou mensuelle. Une augmentation des niveaux de corrélation à l'échelle de la semaine par rapport à l'échelle journalière est néanmoins visible. Elle s'accompagne d'effets significatifs du RG, de REW et de l'interaction entre REW et VPD pour la majorité des génotypes. Parallèlement à ces tendances générales, une variabilité importante entre génotypes a également été mise en évidence dans cette descendance, à la fois sur les paramètres (AMP et ER) et sur leurs corrélations avec les variables environnementales. Cette variabilité est associée à un classement des génotypes différent que ce soit entre paramètres, ou entre les paramètres et leurs corrélations avec les variables environnementales (Figure 6.13). Cette différence suggère un contrôle génétique différent pour ces deux paramètres, ainsi que sur leurs réponses à l'environnement.

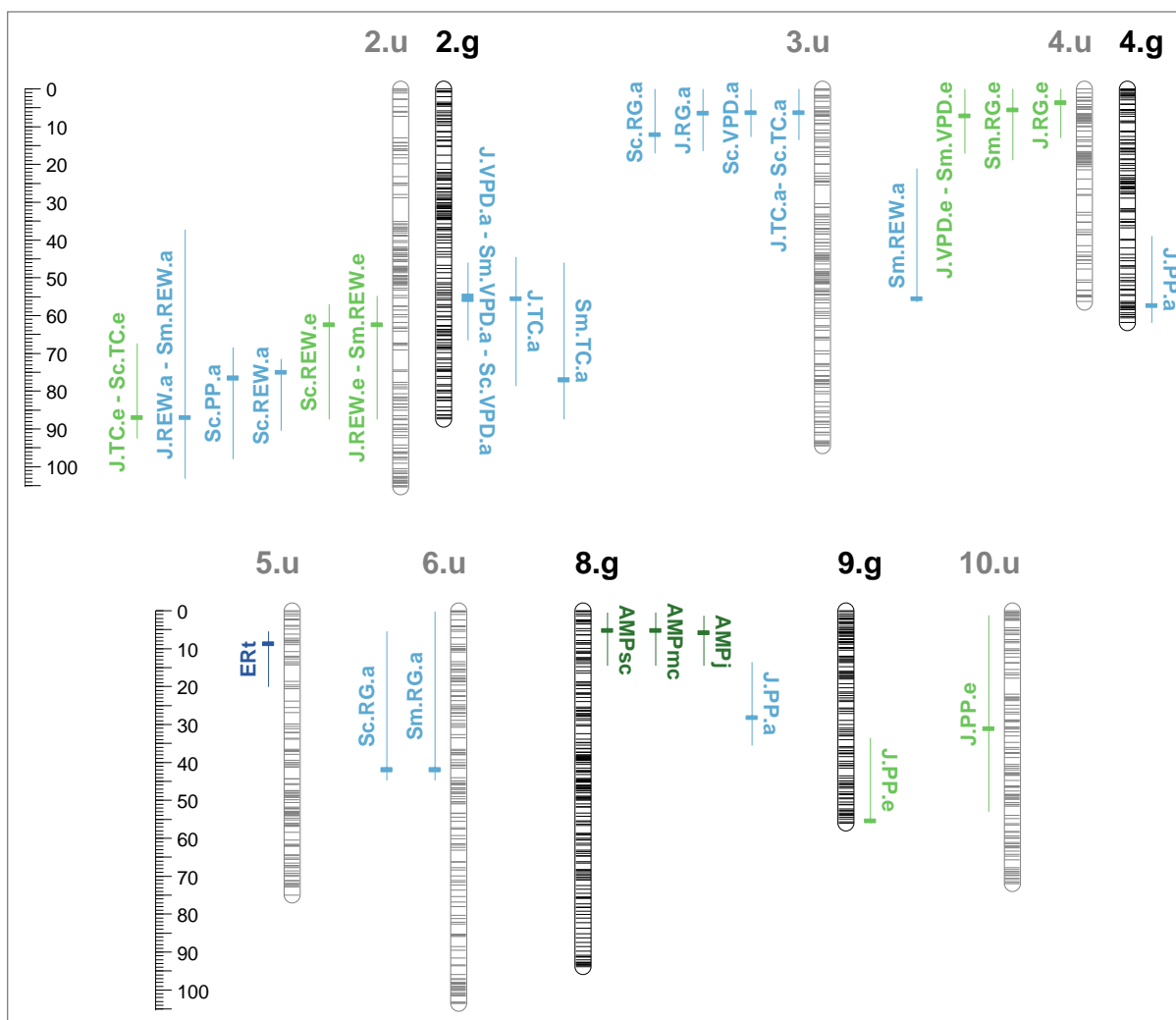


Figure 6.14 : Localisation des QTL détectés pour les paramètres AMP et ER pour les deux parent *E.g* (g) et *E.u* (u). Ces deux paramètres sont calculés à différentes échelles de temps : jour (j) semaine (s) et mois (m) et moyennés sur l'ensemble de la période d'étude. Les corrélations entre ces deux paramètres et les variables environnementales sont également présentées. Les caractères associés à AMP sont présentés en bleu et ceux associés à ER en vert. Le seuil de significativité des QTL est de 5% au niveau du génome.



### 6.4.3 Architecture génétique des micro-variations du rayon

L'architecture génétique des micro-variations du rayon a été analysée sous deux angles différents. Le premier utilise des caractères intégrateurs (AMP et ER moyens) et des indicateurs de la plasticité aux variables environnementales (coefficients de corrélation). Le deuxième compare l'architecture génétique des paramètres journaliers (AMP et ER) sur deux semaines contrastées.

#### 6.4.3.1 Analyse QTL des paramètres moyens et de leur plasticité

Dans cette première approche, les paramètres AMP et ER moyennés sur les deux années d'étude, ainsi que les coefficients de corrélation entre paramètres et variables environnementales ont été utilisés. L'analyse QTL sur ces caractères a permis de mettre en évidence 23 et 11 QTL pour le parent *E.u* et le parent *E.g* respectivement (Figure 6.14, Tableau S6.6). Pour *E.u*, ces QTL sont associés à 21 caractères répartis entre AMP (52%) et ER (48%). À l'inverse, pour *E.g*, sur les 10 caractères présentant des QTL, un seul est associé à ER et les 9 autres sont associés à AMP. Le pourcentage de variance expliqué par un QTL varie entre 3.5% et 7% pour *E.u*, et entre 3.7% et 9.1% pour *E.g*. Les QTL d'*E.u* sont répartis sur 6 LG (2, 3, 4, 5, 6 et 10) et ceux d'*E.g* sur 4 LG (4, 8 et 9). Pour les paramètres moyens, seulement deux régions génomiques ont été mises en évidence : une sur le LG5 à 8.2 cM pour ER (*E.u*) et l'autre sur le LG5 à 8.2 cM pour ER (*E.g*). Pour les corrélations, 9 régions génomiques ont été mises en évidence pour *E.u* et 5 pour *E.g*. Parmi ces 14 régions, 4 sont associées à des corrélations avec plusieurs variables environnementales. La première sur le LG2 (86.5 cM, *E.u*) est associée à des QTL de corrélation entre AMP et TC, et entre ER et REW. La seconde sur le LG2 (54.1-55.3 cM, *E.g*) rassemble des QTL de corrélation entre AMP et TC ou VPD. La troisième est sur le LG3 (5.82 cM, *E.u*) et comporte des QTL de corrélation entre AMP et TC, VPD ou RG. Enfin la quatrième région est sur le LG4 (5.1-6.6 cM, *E.u*) et associe des QTL de corrélation entre ER et TC ou VPD.

#### 6.4.3.2 Analyse QTL des paramètres dans deux environnements contrastés

Dans cette seconde approche, l'architecture génétique des paramètres journaliers AMP et ER est analysée sur deux semaines contrastées au niveau de la disponibilité en eau du sol

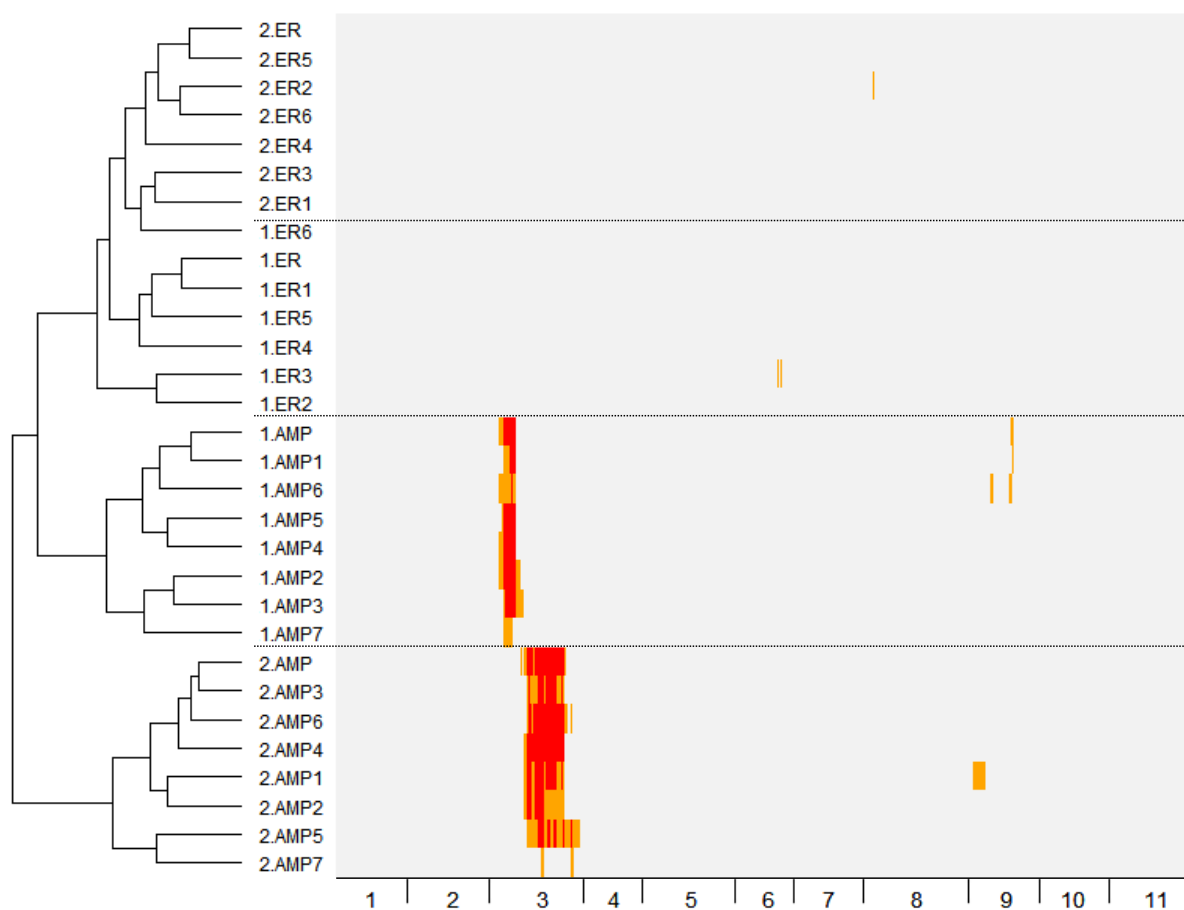


Figure 6.15 : Heatmap des LOD score le long des 11 LG du génome d'*E.u.* L'ensemble des caractères des deux semaines est représenté. La classification hiérarchique des caractères est réalisée sur la base des valeurs de LOD score avec la méthode Ward et une distance euclidienne. En rouge  $\text{LOD} > 3.5$ , en orange  $3.5 > \text{LOD} > 2.5$  et en gris  $\text{LOD} < 2.5$ .





et de la demande évaporatoire. La semaine S1 est caractérisée par une forte demande évaporatoire et une disponibilité en eau importante. La semaine S2 présente une très faible disponibilité et une demande évaporatoire moyenne.

Au total, 20 QTL ont été cartographiés pour le parent *E.u* et 8 pour le parent *E.g* (Tableau S6.7). La majorité des QTL détectés le sont pour le paramètre AMP chez les deux parents avec 95% pour *E.u* et 75% pour *E.g*. Les QTL expliquent en moyenne 7.1% et 6.9% de la variabilité phénotypique, et au maximum 9.4% et 7.8% respectivement pour, *E.u* et *E.g*. Leur répartition sur le génome est assez restreinte puisque seulement 4 LG (3, 6, 8 et 9) sont impliqués pour *E.u* et 2 LG (8 et 10) pour *E.g*. De plus, pour *E.u* deux hotspots localisés sur le LG3 (10.8-12.7 cM et 43.2-45 cM) regroupent 45% et 35% des QTL détectés pour ce parent. De façon similaire, pour *E.g* un hotspot localisé sur le LG8 à 25.8 cM rassemble 75% des QTL détecté pour ce parent. Un effet marqué de l'environnement sur l'architecture génétique d'AMP est visible chez les deux parents. En effet, pour le parent *E.g* aucun QTL n'a été détecté pour S1. Pour *E.u* les deux hotspots situés sur le LG3 sont spécifiques à une seule semaine (Figure 6.15). Cette différence d'architecture génétique entre S1 et S2 indique une interaction QTL  $\times$  environnement forte et donc un effet marqué de la saison sur le contrôle génétique d'AMP.

## 6.5 Discussion

### 6.5.1 La croissance ponctuelle et les paramètres du cycle circadien

Cette étude a permis d'avoir une vision exhaustive de la dynamique de croissance et de sa variabilité au sein de la descendance, grâce d'une part aux mesures ponctuelles sur 960 génotypes, et d'autre part à la caractérisation en continu des micro-variations du rayon sur un sous ensemble de 222 génotypes. Le lien entre les différentes échelles d'analyse de la croissance a donc pu être analysé (Figure S6.8). AMP est corrélée négativement aux mesures ponctuelles de hauteur et de diamètre (au maximum -0.34 pour D22). Pour ER, les corrélations sont positives avec un maximum de 0.51 pour la corrélation entre l'accroissement obtenu à partir des données de micro-dendrométrie et l'accroissement en diamètre sur la même période. Cette différence peut s'expliquer par le fait que la croissance radiale n'est pas uniforme sur toute la circonférence du tronc et notamment à la base du tronc (hauteur où la mesure est effectuée). La mesure de croissance à partir des micro-dendromètres de type ponctuel est sous-estimée par rapport à la mesure ponctuelle de croissance. Bien que cet inconvénient ait



été rapporté [2, 28], peu d'études mettent en lien les deux types de mesures sur autant d'individus. Certaines études mettent en relation la mesure de micro-dendromètres avec la formation de bois (taille des cernes) et montrent une bonne concordance entre les deux [11].

### 6.5.2 Relations entre les paramètres du cycle circadien et les variables environnementales

Les mécanismes physiologiques reliés au cycle circadien sont étudiés depuis plusieurs décennies et sont intrinsèquement liés aux mouvements et au stockage de l'eau dans les tissus élastiques du tronc [9]. Dans cette étude, le choix a été fait de n'utiliser que deux paramètres (AMP et ER) pour caractériser le cycle circadien des micro-variations du rayon. De nombreuses études utilisent l'un et/ou l'autre de ces paramètres pour mettre en lien la croissance, le fonctionnement hydrique de l'arbre et les variables environnementales [6, 11, 12, 29, 30]. Il existe cependant d'autres paramètres, comme la pente de contraction ou la durée des différentes phases du cycle circadien [2, 5, 31]. Une autre approche consiste à retirer la tendance générale de croissance présente dans le signal, ce qui permet ainsi d'obtenir le fonctionnement hydrique "pur" de l'arbre [32]. De fait, toute la variabilité du signal n'a pas été explorée dans cette étude. Néanmoins, ces deux paramètres ont permis de mettre en évidence l'effet des variables environnementales sur le signal de micro-variations. En effet, des corrélations positives fortes à l'échelle journalière entre VPD et AMP ont été montrées dans cette étude, soulignant l'effet prépondérant de la demande évaporatoire sur AMP. Cette corrélation entre VPD et AMP a été montrée chez différentes espèces. Par exemple, Deslauriers *et al.* [11] ont mis évidence des corrélations de 0.48 à 0.69 entre ces deux caractères sur *Larix decidua*, *Pinus cembra* et *Picea abies*. Sur *Pseudotsuga menziesii*, une corrélation de 0.89 est mise en évidence entre VPD et AMP sur la période de végétation [29]. Cette corrélation est attendue du fait du lien étroit entre la transpiration de l'arbre et l'amplitude de contraction journalière [3, 11]. Dans une moindre mesure, la température et le rayonnement (variables corrélées au VPD) ont également été reliés à l'AMP dans cette étude. Cet effet de la température sur AMP a été rapporté dans plusieurs études, avec des proportions variables, selon l'espèce ou la période considérée [11, 12, 29]. Les précipitations et le contenu en eau du sol ne sont quant à eux que très peu reliés à AMP. Bien que ce résultat ne soit pas nécessairement attendu, il a également été rapporté sur plusieurs espèces de région tempérée [11]. Contrairement à AMP, de faibles corrélations entre ER et les variables environnementales à l'échelle journalière ont été mises en évidence dans cette étude. Des résultats plus contrastés ont été rapportés sur *Abies balsamea*, avec des corrélations positives



pour les précipitations et la température maximale, décomposées suivant les phases du cycle circadien [6]. Une approche plus détaillée sur les variables environnementales pourrait donc permettre d'affiner le lien entre ER et l'environnement au niveau journalier. Par ailleurs, le lien entre ER et les variables environnementales est renforcé par le changement d'échelle (semaine, mois), notamment pour TC, RG et PP, corrélés positivement avec ER. Cet effet du climat à plus long terme a également été montré sur *Picea abies* avec une intégration sur cinq jours [7]. Les auteurs ont mis en évidence des corrélations positives avec les précipitations et le contenu en eau du sol, et négatives avec la demande évaporatoire et la température. Des résultats similaires ont également été rapportés par Zweifel *et al.* [33]. Ils montrent une corrélation positive entre l'accroissement en rayon et la pluviométrie sur des périodes de dix jours.

Parallèlement à l'effet de l'environnement sur les paramètres, une variabilité importante entre génotypes a été montrée, pour les relations entre les paramètres du cycle circadien et les variables environnementales. Cette différence a été utilisée pour caractériser la plasticité des génotypes face aux variations de la demande évaporatoire et de la disponibilité en eau du sol. À notre connaissance, cette étude est la seule à présenter la variabilité intraspécifique sur une population aussi large pour ce type de caractère. En effet la plupart des études comparent des espèces différentes [5, 11, 13, 31, 32] ou des conditions environnementales différentes [12, 34, 35] mais ne s'attachent pas à décrire la variabilité interindividuelle comme c'est le cas ici.

### 6.5.3 L'architecture génétique de la croissance et de la plasticité des micro-variations radiales

Dans un premier temps, les résultats sur l'architecture génétique de la croissance totale ont permis de confirmer l'effet de la taille de la population sur la détection de QTL (effet Beavis [36, 37]). Un plus grand nombre de QTL par caractère ainsi que des pourcentages de variance expliquée plus faibles ont été mis en évidence grâce à l'effectif de 960 individus. Ces résultats bien qu'attendus soulignent le déterminisme polygénique de la croissance. Par ailleurs, le lien entre architecture génétique des accroissements au jeune âge et celle de la croissance totale à l'âge de rotation, mise en évidence dans le chapitre 3, semble se confirmer dans ce chapitre sur les trois premières années de croissance.





Dans un second temps, l'analyse des variations journalières du rayon, en relation avec les variables environnementales, nous a permis de déterminer les régions génomiques impliquées dans le contrôle génétique de la plasticité des paramètres du cycle circadien au sein de la descendance hybride *E. urophylla* x *E. grandis*. Plusieurs QTL ont été mis en évidence pour ces caractères de réponse à l'environnement, montrant ainsi un contrôle génétique de l'interaction G×E. Ce contrôle génétique s'est révélé être très différent entre AMP et ER, avec seulement un QTL commun entre les deux paramètres. On retrouve donc au niveau génétique la différence de plasticité entre AMP et ER mise en évidence au niveau phénotypique. Ceci renforce l'idée que bien ces deux paramètres soient issus du même signal de micro-variations, ils traduisent un fonctionnement différent. De plus, que ce soit pour AMP ou pour ER, les régions génomiques associées à la demande évaporatoire (VPD, RG, TC) sont distinctes de celles associées à la disponibilité en eau (REW, PP). Ce résultat suggère un contrôle génétique différent entre la réponse à la demande évaporatoire au niveau du système aérien et celle à la disponibilité en eau au niveau du système racinaire. Chez les arbres forestiers, aucune étude QTL n'a analysé à ce jour l'architecture génétique de la réponse à ces deux types de contraintes. Chez le maïs, un déterminisme génétique commun entre ces deux types de réponse a été mis en évidence grâce à l'étude de la plasticité de la croissance foliaire [38]. Bien que le matériel végétal soit très différent (pérenne vs. annuelle), ces résultats montrent qu'en fonction du niveau d'intégration (organe, plante entière), les conclusions peuvent être différentes. En effet, cette étude est centrée sur un seul organe (feuille), dont la croissance est sensible aux deux types de contrainte. Dans notre cas, la micro-variation du tronc intègre différents processus liés au fonctionnement hydrique de l'arbre (transpiration au niveau foliaire, absorption racinaire,...). Il serait intéressant d'analyser le contenu en gène de ces régions, afin de déterminer si les gènes potentiellement impliqués sont associés à des fonctions distinctes ou similaires.

Par ailleurs, l'origine de la variabilité génétique fonctionnelle, au sein de ce croisement, semble différente en fonction du parent considéré. En effet, un seul QTL pour ER a été détecté pour le parent *E.g* et un nombre plus faible de régions a été mis en évidence pour celui-ci. Même s'il est difficile d'extrapoler ces résultats au niveau des deux espèces (*E.u* et *E.g*), la différence d'adaptation aux conditions congolaises [39], va dans le sens d'une réponse contrôlée principalement par la variabilité allélique du parent *E.u*.

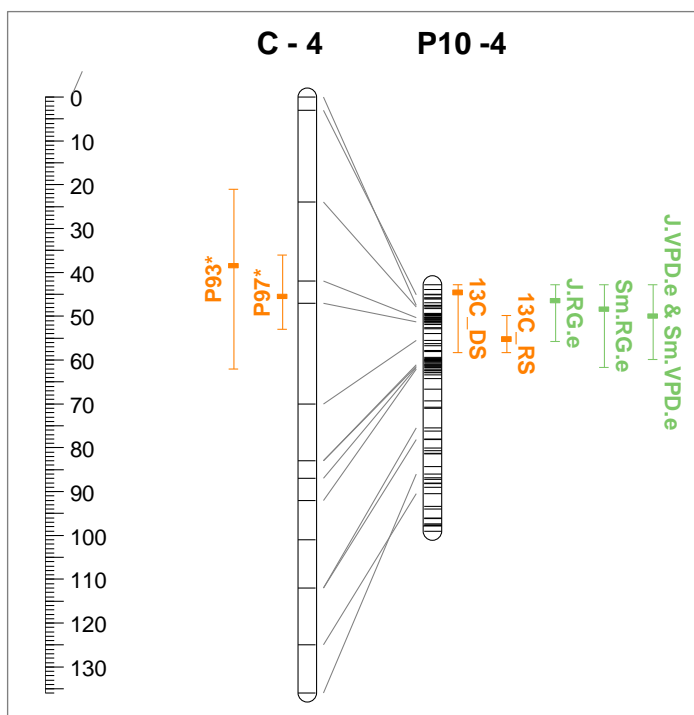


Figure 6.16 : Localisation des QTL de  $\delta^{13}C$  (en orange) et des QTL associés à ER (en vert) sur le LG4 de *E.u.* pour la carte consensus (C) entre P93 et P97 et la carte génétique du dispositif P10.

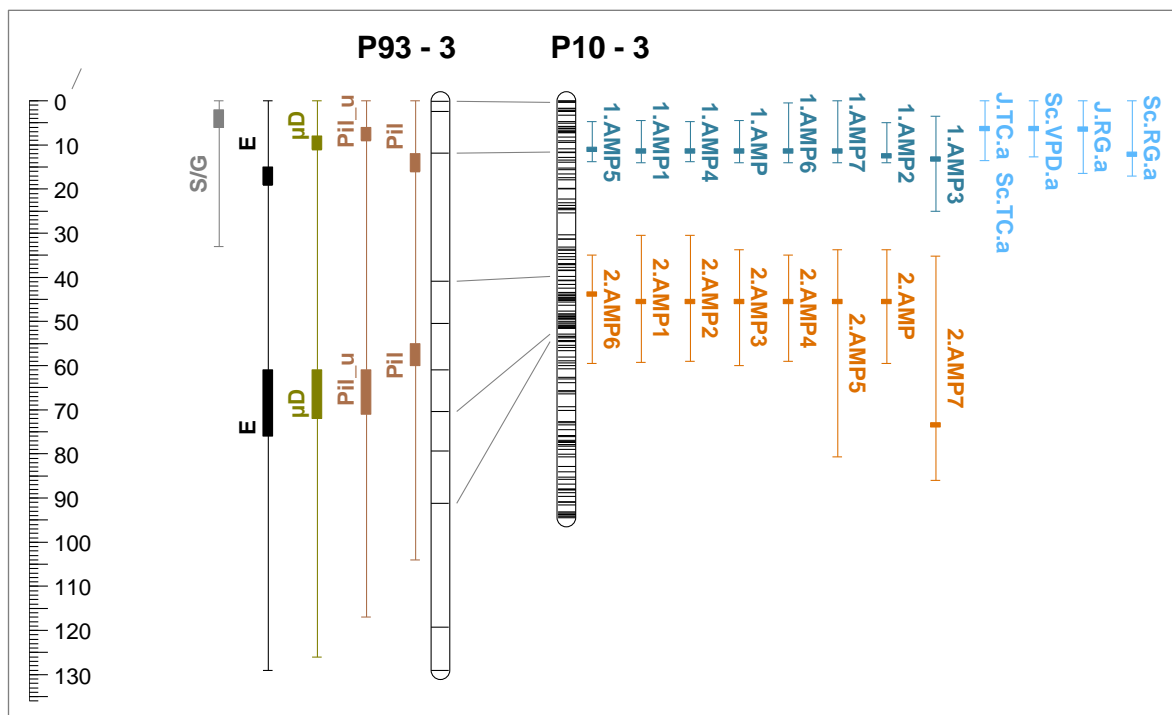


Figure 6.17 : Colocalisation entre des QTL de propriétés du bois (dispositif P93) et des paramètres de micro-variations du rayon (dispositif P10) sur le LG3 du parent *E.u.* Pour le dispositif P10, les QTL associés à l'amplitude en saison des pluies (S1) sont en bleu foncé, ceux associés à l'amplitude en saison sèche (S2) sont en orange, et ceux associés aux corrélations entre l'amplitude et les variables environnementales sont en bleu clair. Pour le dispositif P93, les données sont issues de [40]. Les QTL pour le Pilodyn (Pil et Pil\_u = sous écorce) sont en marron, ceux pour la micro-densitométrie moyenne ( $\mu D$ ) sont en kaki, ceux pour le module d'élasticité sont en noir et celui pour le rapport S/G en gris.



#### 6.5.4 Lien entre l'architecture génétique des micro-variations et celle du $\delta^{13}\text{C}$ ainsi que celle des propriétés du bois

L'analyse des colocalisations entre les caractères issus du signal journalier et d'autres traits, comme le  $\delta^{13}\text{C}$  (chapitre 5) et les propriétés du bois [40], a révélé la richesse de l'information contenue dans ce signal journalier en termes de variabilité génétique. Trois colocalisations intéressantes ont été mises en évidence.

La première colocalisation, entre le  $\delta^{13}\text{C}$  et la réponse d'ER au VPD et à RG est située sur le LG4 d'*E.u* (Figure 6.16). L'optimum de LOD du QTL de  $\delta^{13}\text{C}$  en saison sèche est étroitement associé à celui de deux QTL de réponse d'ER à RG, à l'échelle de la journée et de la semaine (J.RG.e et Sm.JG.e). Comme montré par l'analyse d'enrichissement GO (chapitre 5), les QTL de  $\delta^{13}\text{C}$  de saison sèche semblent traduire une réponse des arbres au stress durant la première saison sèche. Cette colocalisation avec un QTL de réponse d'ER avec le rayonnement global est intéressante, car le rayonnement est plus faible en saison sèche (cf. §6.4.1). Ce résultat est en phase avec les corrélations observées entre diamètre et  $\delta^{13}\text{C}$  de saison sèche. Cette réponse au stress de saison sèche pourrait être une réponse combinée, entre une contrainte hydrique marquée à cette saison et un rayonnement incident limitant, révélant les différences entre génotypes. Même s'il est difficile de conclure sur ces colocalisations, une liaison étroite entre gènes semble être l'hypothèse la plus probable au regard des optimums de LOD pour ces QTL.

Deux autres colocalisations situées sur le LG3 (*E.u*, Figure 6.17) apportent des éléments de compréhension supplémentaires sur la variabilité génétique du signal journalier. En effet, les deux régions de QTL majeures pour les propriétés du bois (notamment la micro-densitométrie moyenne sur le rayon à cinq ans) qui ont été mises en évidence sur ce LG [40], semblent correspondre à deux régions QTL pour l'amplitude journalière. De plus, ces régions QTL pour l'amplitude sont associées à deux semaines contrastées au niveau du climat (saison sèche vs saison des pluies). Un déterminisme génétique commun entre les propriétés du bois et l'amplitude de contraction journalière du tronc est cohérent avec le fonctionnement physiologique de l'arbre. En effet, les micro-variations du rayon ont été mises en relation avec les propriétés du bois, comme la densité ou l'angle des microfibrilles [34, 41]. Par ailleurs ces études sur *E. nitens* montrent que la réduction de la disponibilité en eau s'accompagne d'une réduction de la croissance et d'une augmentation de la densité du bois. Ces deux colocalisations illustrent que la réponse de l'arbre à la contrainte hydrique aurait un déterminisme génétique spécifique, avec une incidence sur les propriétés du bois. La





variabilité génétique associée à l'amplitude de contraction journalière observée au jeune âge, caractère lié à l'élasticité des tissus périphériques (comme le phloème ou le cambium) [42], pourrait expliquer à plus long terme les différences génétiques des propriétés mécaniques du bois au stade adulte.

## Figures et tableaux additionnels

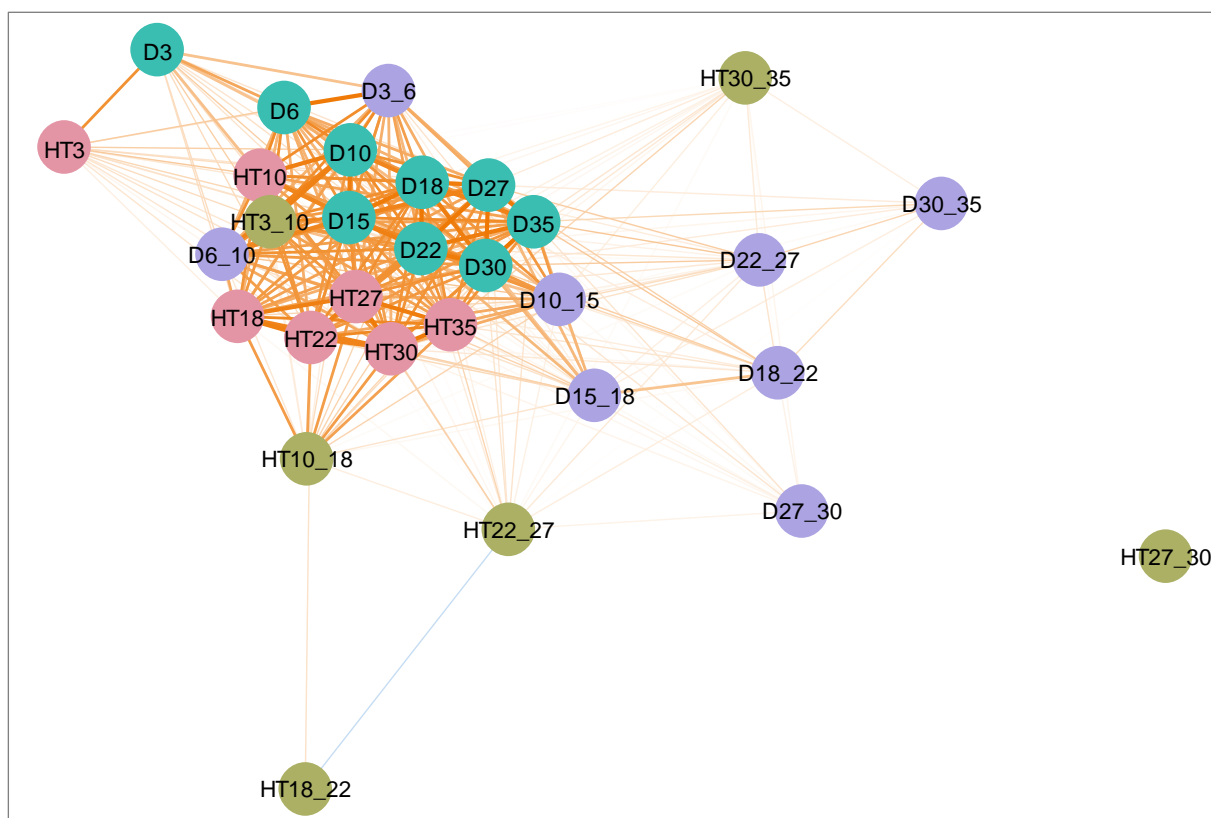


Figure S6.1: Diagramme des corrélations (Spearman) pour les mesures ponctuelles (D, HT) et les accroissements associés. La distance entre les cercles représente la structuration des corrélations entre variables. L'intensité et la couleur des traits entre les cercles indiquent la valeur de la corrélation : en orange les corrélations positives, en bleu les corrélations négatives, et plus le trait est foncé plus la corrélation est forte. Les cercles sont en bleu pour D, en rouge pour HT, pour les accroissements en violet pour D et en vert pour HT.

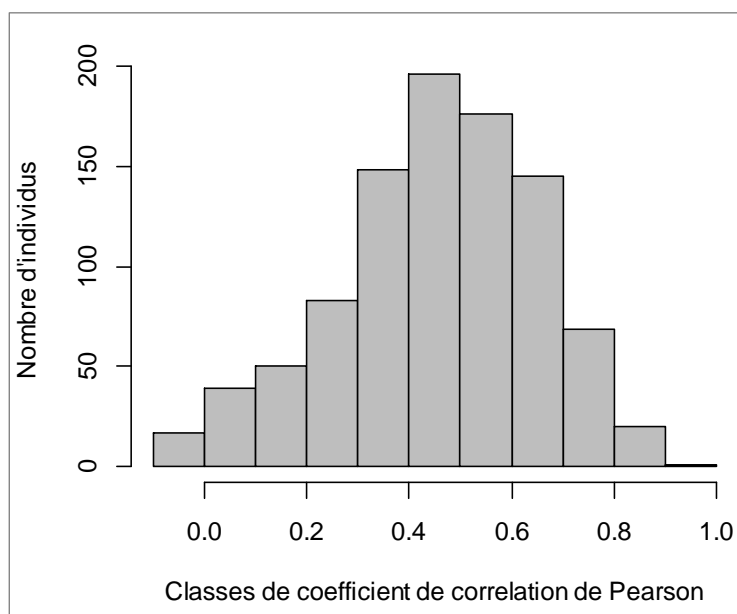


Figure S6.2: Distribution des coefficients de corrélation entre les accroissements en diamètre et l' $I_{DM}$  pour l'ensemble des génotypes.

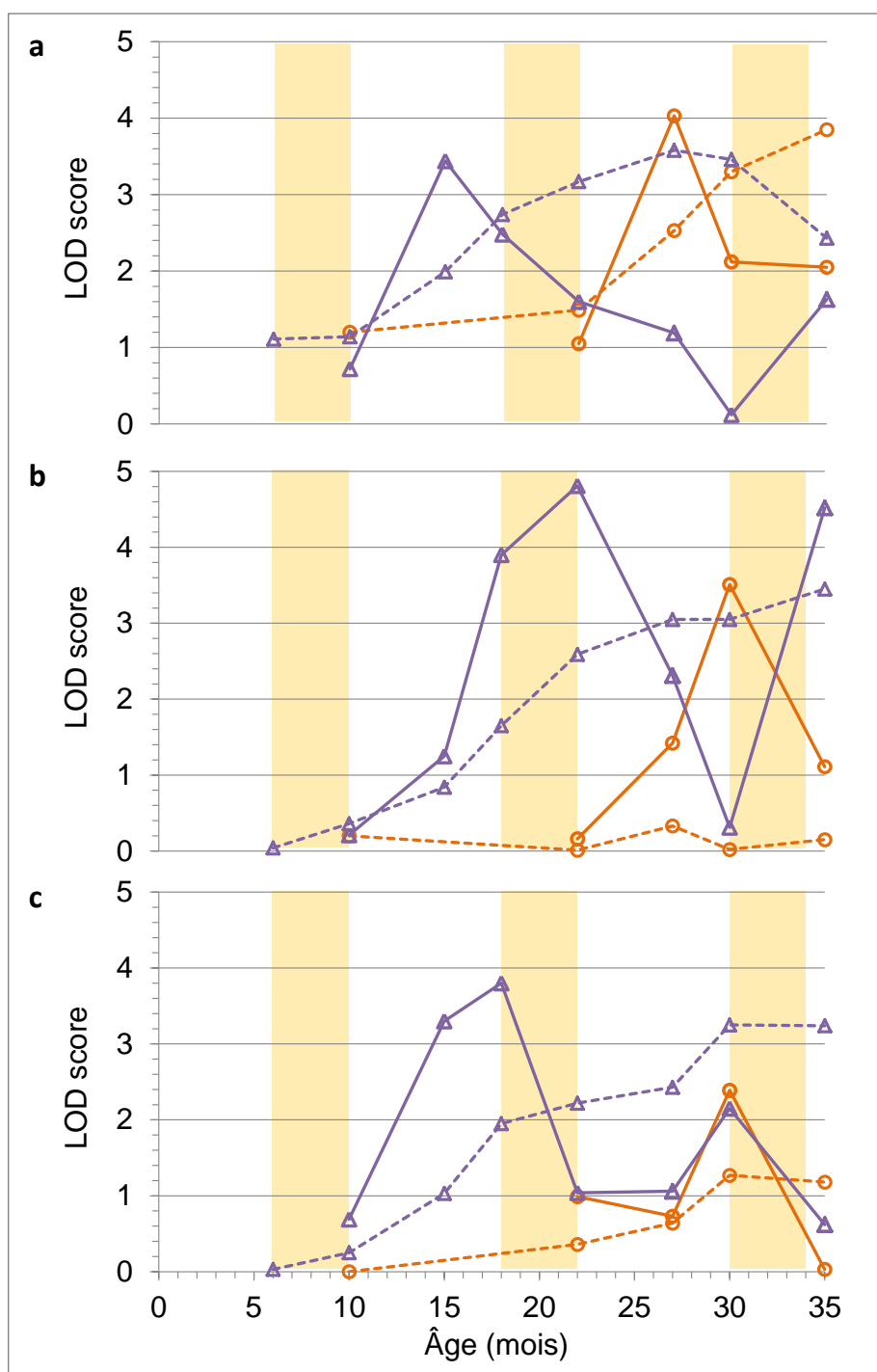


Figure S6.3 : Evolution du LOD score au cours du temps au niveau de trois régions QTL. (a) Le LG4 à 10.3 cM et (b) le LG8 à 40.3 cM pour *E.u.* (c) le LG8 à 40.3 cM pour *E.u.* HT est indiquée en orange avec des cercles et D en violet avec des triangles. La croissance cumulée est représentée en traits pointillés et les accroissements en traits pleins. Les bandes verticales orange indiquent les saisons sèches.

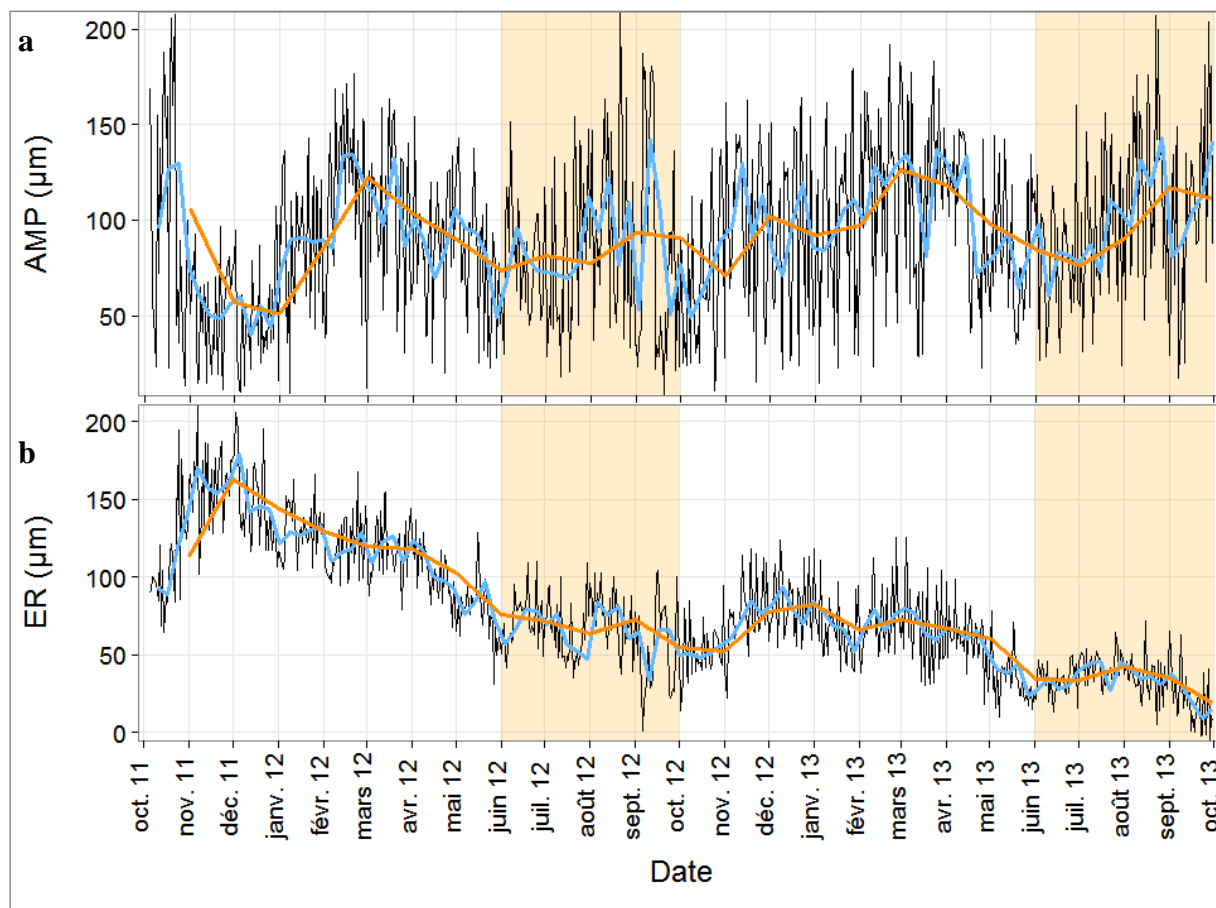


Figure S6.4: Evolution de l'amplitude journalière (AMP, a) et de l'accroissement net du rayon (ER, b). La moyenne sur l'ensemble des arbres suivis est représentée par jour (trait noir), par semaine (trait bleu) et par mois (trait orange). Les bandes verticales orange indiquent les saisons sèches.

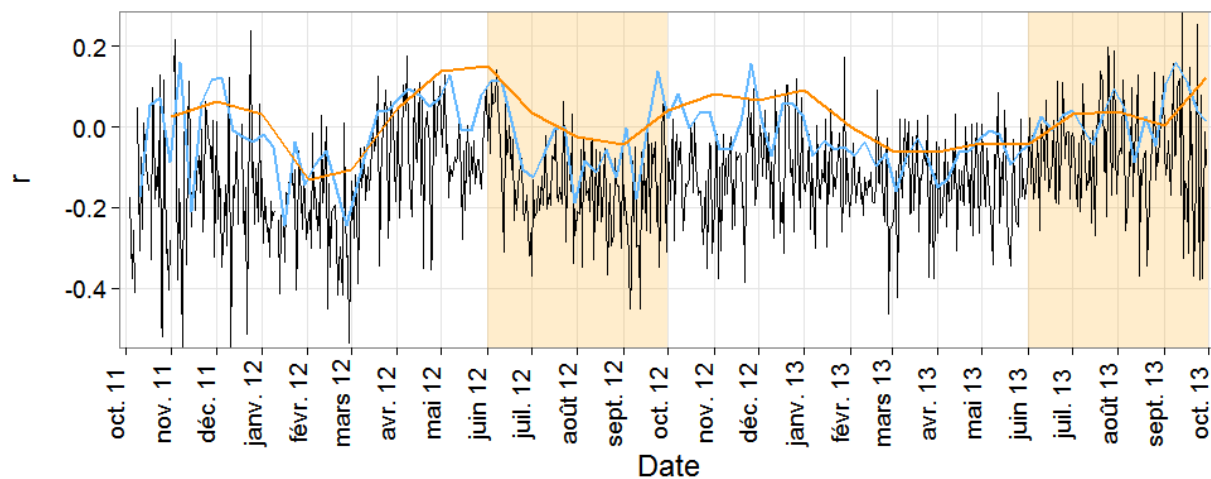


Figure S6.5 : Evolution des corrélations entre AMP et ER sur l'ensemble des génotypes par jour (trait noir), par semaine (trait bleu) et par mois (trait orange). Les bandes verticales orange indiquent les saisons sèches.

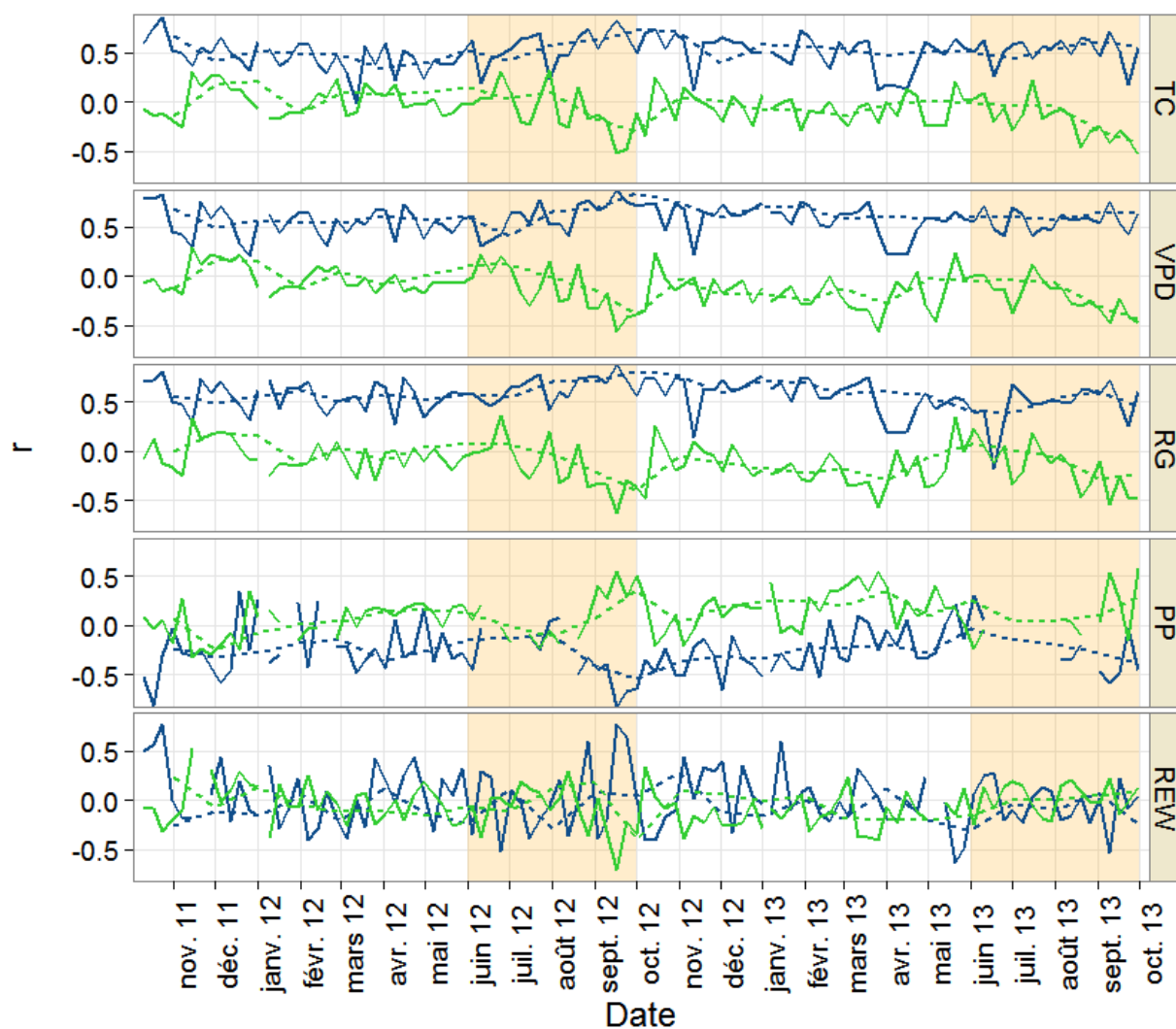


Figure S6.6: Evolution du coefficient de corrélation de Pearson ( $r$ ) entre les paramètres journaliers (AMP en bleu ou ER en vert) et les variables environnementales journalières. Les corrélations sont calculées sur la semaine (trait plein) ou le mois (trait pointillé) sur l'ensemble des génotypes. La valeur journalière est le maximum pour TC, la moyenne pour VPD, RG et REW, et la somme pour PP.

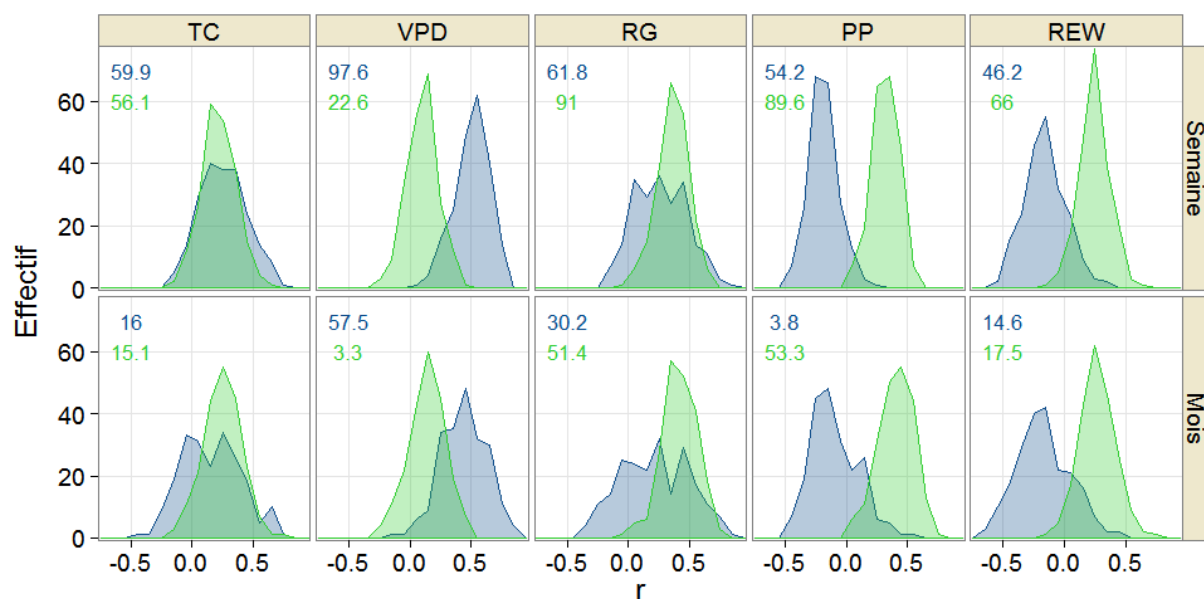


Figure S6.7 : Distribution des coefficients de corrélation de Pearson ( $r$ ) entre les paramètres (AMP en bleu ou ER en vert) et les variables environnementales sur l'ensemble de la période. Les corrélations sont calculées pour chaque génotype sur toute la durée de l'expérimentation, avec les données journalières cumulées à la semaine ou au mois. La valeur journalière est le maximum pour TC, la moyenne pour VPD, RG et REW, et la somme pour PP. Les chiffres indiquent le pourcentage de génotypes présentant une corrélation significative.

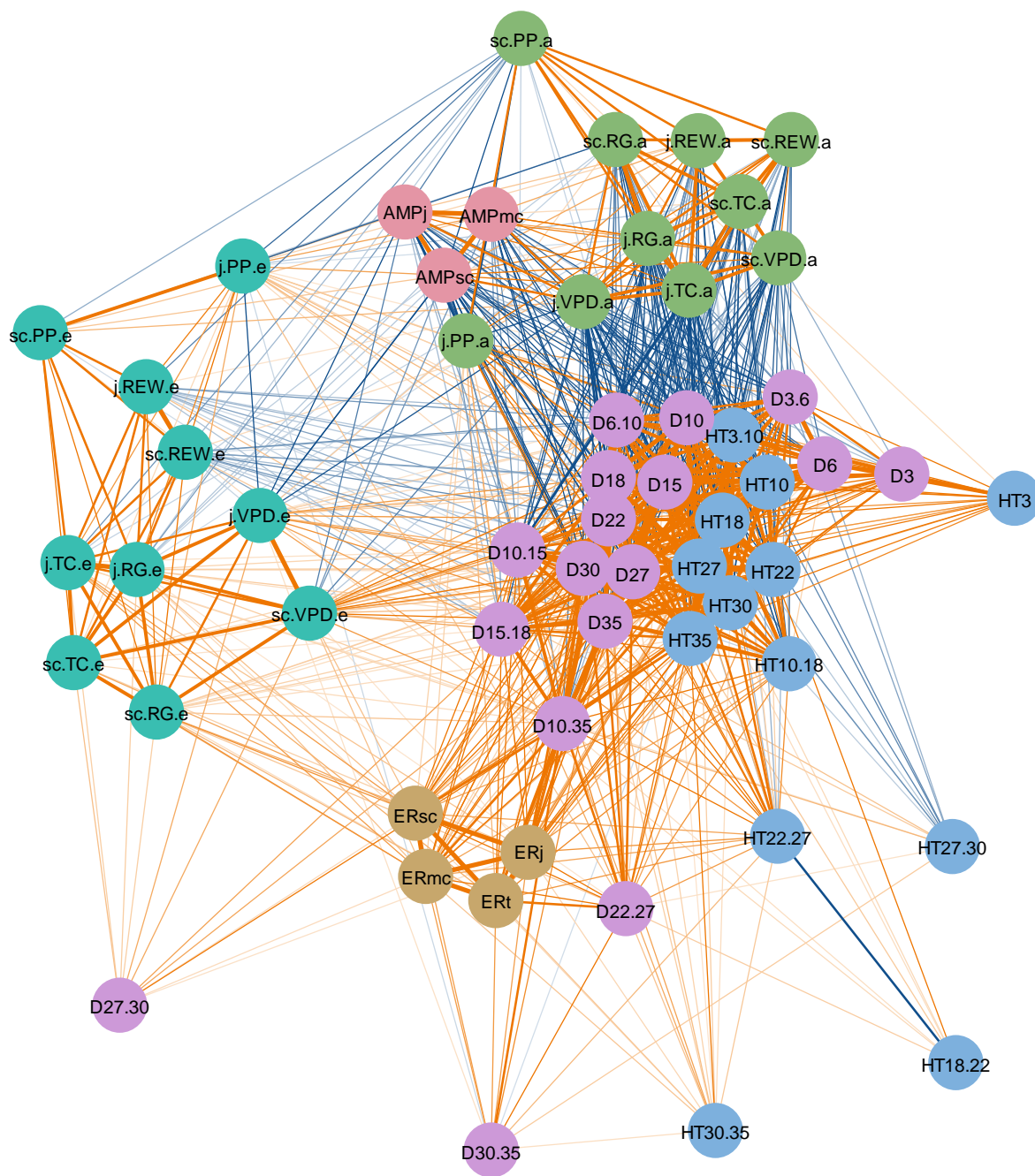


Figure S6.8 : Diagramme des corrélations (Spearman) pour les caractères de micro-variations (AMP et ER) et de croissance ponctuelle (HT et D). La distance entre les cercles représente la structuration des corrélations entre variables. L'intensité et la couleur des traits entre les cercles indiquent la valeur de la corrélation : en orange les corrélations positives, en bleu les corrélations négatives, plus le trait est foncé plus la corrélation est forte. Les cercles roses représentent AMP moyenne sur l'ensemble de la période d'étude pour les données journalières (j) et pour les cumuls sur la semaine (sc) et le mois (mc). De la même manière ER est représenté en marron avec le cumul total (t). En turquoise les corrélations entre AMP (a) et les variables environnementales à l'échelle de la journée (J) et de la semaine (Sc, Sm). Le même code est utilisé pour ER (e) représenté en vert. HT est représentée en violet et D en bleu. Seules les corrélations significatives sont représentées.





Tableau S6.1 : Statistiques descriptives des caractères de croissance ponctuelle (HT: hauteur, D: diamètre). Avec le nombre de valeurs (N), minimum (Min), maximum (Max), moyenne, écart-type (SD) et le coefficient de variation phénotypique (CVp).

Traits	N	Min	Max	Mean	SD	CVp
HT3	924	0.1	1.12	0.6	0.15	0.27
HT10	960	0.5	5.5	2.5	0.95	0.37
HT18	221	5	12.3	9.1	1.33	0.15
HT22	960	5.2	15.3	10.6	1.7	0.16
HT27	952	6.5	18.7	14	1.95	0.14
HT30	957	7.7	21	16	2.02	0.13
HT35	957	7.9	22.6	17.9	2.26	0.13
HT3_10	924	0.00	0.66	0.27	0.12	0.44
HT10_18	221	0.36	0.91	0.72	0.09	0.13
HT10_22	960	0.31	0.85	0.63	0.09	0.14
HT18_22	221	0.20	0.85	0.51	0.11	0.22
HT22_27	952	0.14	1.24	0.69	0.14	0.2
HT22_30	957	0.18	1.03	0.69	0.1	0.15
HT27_30	951	0.13	1.20	0.69	0.14	0.21
HT30_35	957	0.00	0.67	0.36	0.1	0.26
D3	600	1.1	17.8	5	2.1	0.42
D6	329	2.5	38.9	18	7.4	0.41
D10	959	5	80	33	14.45	0.43
D15	960	15	116	66	19.58	0.29
D18	960	27	155	90	22.51	0.25
D22	959	38	174	108	24.23	0.22
D27	958	48	192	128	26.6	0.21
D30	957	54	204	138	27.8	0.2
D35	957	57	226	148	29.26	0.2
D3_6	317	0	10.6	4.56	2.07	0.45
D3_10	599	0.5	9.1	4.23	1.85	0.44
D6_10	329	0.5	11.5	4.18	2.05	0.49
D10_15	959	1.8	10.1	6.59	1.46	0.22
D10_18	959	2.3	9.9	6.52	1.29	0.2
D10_22	958	2.3	9	5.85	1.09	0.19
D15_18	960	0.4	12.1	6.42	1.62	0.25
D18_22	959	0.4	9.7	4.45	1.06	0.24
D22_27	957	0.3	9.1	4.04	1.06	0.26
D22_30	956	0.4	7.2	3.84	0.88	0.23
D27_30	957	0.2	9	3.51	1.26	0.36
D30_35	957	0	5.4	1.98	0.73	0.37



Tableau S6.2 : Matrice des coefficients de corrélation de Spearman pour les caractères de croissance ponctuelle. Seules les corrélations significatives sont indiquées

	HT3	HT10	HT18	HT22	HT27	HT30	HT35	HT3_10	HT10_18	HT10_22	HT18_22	HT22_27	HT22_30	HT27_30	HT30_35		D3	D6	D10	D15	D18	D22	D27	D30	D35	D3_6	D3_10	D6_10	D10_15	D10_18	D10_22	D15_18	D18_22	D22_27	D22_30	D27_30		
HT10	.51																																					
HT18	.29	.81																																				
HT22	.35	.81	.94																																			
HT27	.36	.80	.94	.94																																		
HT30	.35	.76	.92	.92	.97																																	
HT35	.34	.75	.90	.90	.95	.97																																
HT3_10	.40	.99	.80	.81	.80	.76	.74																															
HT10_18		.31	.78	.75	.77	.77	.75	.30																														
HT10_22	.08	.37	.66	.82	.74	.75	.74	.38	.86																													
HT18_22				.37	.16	.16	.16			.56																												
HT22_27	.15	.21	.24	.16	.46	.43	.44	.21	.30	.07	-.40																											
HT22_30	.11	.12	.14	.10	.35	.44	.44	.11	.23	.08	-.35	.82																										
HT27_30		-.08				.16	.14	-.08		.10			.49																									
HT30_35	.11	.21		.24	.27	.26	.44	.20		.22		.23	.18																									
D3	.84	.57	.39	.40	.39	.38	.36	.48		.11																												
D6	.50	.88	.63	.67	.67	.64	.63	.85	.18	.24		.19	.11		.17	.71																						
D10	.48	.95	.76	.80	.78	.75	.74	.94	.29	.40		.22	.14		.21	.55	.88																					
D15	.43	.91	.79	.84	.83	.81	.80	.90	.40	.50		.25	.18		.25	.48	.81	.96																				
D18	.39	.86	.77	.85	.85	.83	.82	.85	.44	.56		.28	.21		.27	.43	.76	.91	.97																			
D22	.37	.81	.76	.84	.84	.83	.82	.81	.45	.57		.31	.25		.28	.38	.72	.87	.94	.99																		
D27	.36	.78	.73	.82	.84	.83	.83	.78	.46	.58		.34	.29		.31	.36	.69	.84	.92	.96	.98																	
D30	.35	.77	.71	.81	.83	.83	.83	.76	.44	.58		.35	.31		.33	.35	.68	.82	.90	.95	.97	.99																
D35	.34	.74	.67	.78	.81	.81	.82	.73	.40	.56		.36	.32		.34	.33	.65	.79	.87	.93	.96	.98	.99															
D3_6	.35	.87	.62	.66	.67	.64	.62	.86	.17	.24		.19			.18	.56	.98	.87	.81	.77	.73	.69	.68	.65														
D3_10	.32	.93	.74	.80	.77	.75	.72	.93	.28	.39		.18	.09		.16	.45	.85	.99	.95	.91	.87	.84	.81	.78	.86													
D6_10	.26	.83	.70	.76	.73	.70	.68	.84	.34	.42		.19			.15	.40	.64	.92	.91	.87	.84	.81	.79	.76	.65	.94												
D10_15	.21	.56	.44	.69	.70	.68	.69	.56	.45	.59		.28	.22		.30	.16	.36	.59	.78	.81	.81	.80	.80	.78	.36	.52	.54											
D10_18	.18	.50	.51	.68	.70	.70	.69	.51	.46	.62	.13	.32	.28		.28	.14	.38	.53	.70	.82	.85	.85	.85	.84	.39	.52	.54	.88										
D10_22	.16	.44	.46	.63	.66	.67	.67	.44	.44	.61	.14	.34	.33	.11	.29	.09	.33	.47	.63	.76	.83	.84	.84	.85	.35	.47	.49	.81	.96									
D15_18	.07	.24	.42	.42	.45	.47	.46	.25	.36	.45		.26	.26	.11	.18		.28	.26	.36	.56	.61	.61	.62	.63	.29	.34	.37	.44	.80	.82								
D18_22		.09	.17	.25	.31	.34	.35	.09	.24	.33		.30	.34	.17	.22		.12	.20	.31	.45	.47	.49	.51		.12	.15	.31	.47	.68	.55								
D22_27	.11	.17	.18	.27	.34	.37	.40	.17	.23	.30	.16	.32	.38	.19	.32		.17	.22	.28	.31	.34	.49	.50	.53	.15	.22	.23	.33	.37	.40	.29	.36						
D22_30	.10	.19	.19	.30	.38	.41	.44	.19	.20	.33	.15	.38	.43	.20	.36		.15	.22	.29	.34	.37	.49	.56	.58	.13	.17	.18	.36	.42	.45	.35	.40	.84					
D27_30		.13		.20	.25	.27	.29	.12		.22		.27	.31	.13	.26		.12	.16	.20	.23	.24	.35	.35					.24	.28	.30	.24	.27	.20	.66				
D30_35	.09	.10		.11	.15	.16	.20	.08		.12		.21	.24	.11	.29		.12	.16	.21	.26	.31	.31	.42		.10			.23	.30	.35	.30	.39	.44	.42	.20			



Tableau S6.3 : QTL détectés pour les caractères de croissance ponctuelle.

Parent	Caractère	n	LG	Position (cM)	LOD	BCI	Modèle QTL	PEV	PEV tot	D
<i>E.g</i>	HT22	960	4	8.47	3.04	0 - 45.5	1	1.28	-	0.38
	HT10_22	960	2	7.89	3.48	0 - 64	1	1.47	-	0.02
	D10	959	5	50.75	3.15	42 - 64.5	1	1.19	-	-3.16
	D18	960	2	42.04	3.92	30.5 - 57.8	2	1.34	2.69	5.31
	D18	960	3	78.84	3.35	68.5 - 87	2	1.29	2.69	-5.12
	D22	959	2	42.04	5.2	31 - 47.5	3	2.07	4.87	7.14
	D22	959	3	78.84	3.97	68.5 - 85	3	1.58	4.87	-6.09
	D22	959	5	50.45	2.82	40.3 - 66	3	1.41	4.87	-5.79
	D27	958	2	42.04	5.86	34.3 - 48	3	2.43	5.47	8.5
	D27	958	3	78.84	4.49	68.8 - 84.2	3	1.86	5.47	-7.27
	D27	958	8	5.32	2.96	0 - 69.7	3	1.5	5.47	-6.57
	D30	957	2	42.04	4.58	27.9 - 53.1	3	2.18	5.44	8.4
	D30	957	3	78.84	4.47	68.5 - 83.8	3	1.87	5.44	-7.61
	D30	957	8	5.32	3.83	0.8 - 23.5	3	1.76	5.44	-7.41
	D35	957	2	42.04	5.04	31.1 - 49	3	2.42	5.62	9.32
	D35	957	3	74.51	4.37	68.1 - 82	3	1.79	5.62	-7.85
	D35	957	8	5.32	4	0.5 - 22.5	3	1.87	5.62	-8.06
	D10_15	959	2	31.7	2.78	3 - 71.2	4	1.22	5.12	0.33
	D10_15	959	3	78.55	2.85	65.5 - 87	4	1.24	5.12	-0.33
	D10_15	959	8	15.87	3.3	2.4 - 35	4	1.55	5.12	-0.36
	D10_15	959	10	41.62	3.29	17 - 61.5	4	1.42	5.12	-0.35
	D10_18	959	2	34.55	3.32	2 - 68.5	4	1.45	6.78	0.32
	D10_18	959	3	78.84	4.57	68 - 85	4	2.02	6.78	-0.37
	D10_18	959	8	15.87	4.76	4.5 - 24.5	4	2.22	6.78	-0.39
	D10_18	959	10	20.72	3.46	8.5 - 59.6	4	1.59	6.78	-0.33
	D10_22	958	2	34.55	4.75	27 - 56.5	3	2.18	6.02	0.33
	D10_22	958	3	78.84	4.82	67 - 81.5	3	2.05	6.02	-0.31
	D10_22	958	8	15.87	4.31	4.5 - 24.5	3	2.03	6.02	-0.31
	D15_18	960	3	78.84	3.59	67.5 - 86.7	2	1.45	3.14	-0.39
	D15_18	960	8	13.5	4.18	3 - 23.9	2	1.78	3.14	-0.43
	D18_22	959	2	31.11	7.66	27 - 36.5	2	3.34	4.91	0.4
	D18_22	959	3	74.51	3.22	64.5 - 81	2	1.44	4.91	-0.26
	D22_27	957	3	74.51	3.2	32 - 87	2	1.5	3.09	-0.26
	D22_27	957	7	14.75	3.42	2 - 24.5	2	1.47	3.09	0.26
	D22_30	956	5	14.79	3.09	0 - 25	4	1.31	6.64	0.2
	D22_30	956	6	1.97	3.3	0 - 13.5	4	1.45	6.64	-0.21
	D22_30	956	7	14.75	4.34	3.6 - 27	4	2.01	6.64	0.25
	D22_30	956	8	91.3	3.17	0.8 - 94	4	1.46	6.64	0.21
	D27_30	957	8	9.96	3.18	0 - 93.5	1	1.39	-	-0.3
	D30_35	957	7	58	3	47 - 69.7	2	1.4	3.2	0.17
	D30_35	957	10	38.86	3.61	31.4 - 52.5	2	1.68	3.2	-0.19
<i>E.u</i>	HT3	924	10	10.46	4.65	5.3 - 24	1	1.84	-	0.04
	HT10	960	4	51.35	2.94	0 - 56.5	1	1.31	-	-0.22
	HT22	960	4	52.5	4.5	50 - 55.5	1	2.06	-	-0.49
	HT27	952	4	52	5.45	49 - 55.5	1	2.46	-	-0.62
	HT30	957	4	51	6.16	32 - 55.1	2	2.84	4.12	-0.68
	HT30	957	5	45.57	3.28	29.5 - 62	2	1.45	4.12	-0.49
	HT35	957	4	51	5.83	7 - 55.5	2	2.5	3.96	-0.72
	HT35	957	5	45.77	3.24	33 - 67	2	1.63	3.96	-0.58
	HT10_22	960	4	54.5	3.89	47 - 56.5	1	1.64	-	-0.02
	HT22_27	952	4	28.26	4.86	6.5 - 34.5	3	2.2	5.29	-0.04
	HT22_27	952	5	33.54	3.76	25 - 66.5	3	1.94	5.29	-0.04
	HT22_27	952	8	54.46	3.39	41.8 - 103.7	3	1.63	5.29	-0.04
	HT22_30	957	2	101.23	3.15	78.3 - 105.5	3	1.44	7.04	0.02
	HT22_30	957	4	29.5	7.61	13.7 - 34	3	3.31	7.04	-0.04
	HT22_30	957	5	63.5	5.62	32 - 67	3	2.64	7.04	-0.03



Parent	Caractère	n	LG	Position (cM)	LOD	BCI	Modèle QTL	PEV	PEV tot	D
<i>E.u</i>	HT27_30	951	4	31	3.11	3.2 - 44.3	3	1.41	4.35	-0.03
	HT27_30	951	5	46	3.02	31 - 74.8	3	1.39	4.35	-0.03
	HT27_30	951	8	41.26	3.51	30.9 - 50.5	3	1.59	4.35	0.04
	HT30_35	957	2	98.27	3.06	79.5 - 104.3	4	1.37	7.54	0.02
	HT30_35	957	3	76	4.1	55 - 82.7	4	1.97	7.54	-0.03
	HT30_35	957	5	66.42	5.46	62 - 74.8	4	2.42	7.54	-0.03
	HT30_35	957	10	0.49	4.47	0 - 26	4	1.92	7.54	0.03
	D10	959	4	52	3.39	44.3 - 56.5	1	1.58	-	-3.66
	D15	960	4	52	3.63	5.1 - 56.5	1	1.71	-	-5.17
	D18	960	4	52	3.62	4.8 - 56.5	1	1.78	-	-6.06
	D22	959	4	52	3.39	1.1 - 56.5	1	1.69	-	-6.35
	D27	958	3	53.58	2.95	44.3 - 82	3	1.3	4.57	-6.07
	D27	958	4	10.33	3.58	4 - 55	3	1.65	4.57	-6.84
	D27	958	8	40.37	3.05	0 - 59	3	1.43	4.57	6.36
	D30	957	4	10.33	3.46	3.4 - 54.7	2	1.63	3.24	-7.11
	D30	957	8	40.37	3.05	13 - 67.9	2	1.58	3.24	6.99
	D35	957	8	40.37	3.45	13.5 - 47.5	1	1.93	-	8.13
	D10_15	959	4	7.1	3.62	0 - 31	1	1.5	-	-0.36
	D10_18	959	4	10.33	4.38	1 - 30.5	2	1.94	3.6	-0.36
	D10_18	959	8	40.37	3.52	28.4 - 75.5	2	1.63	3.6	0.33
	D10_22	958	4	10.33	3.83	0.5 - 31	2	2.03	4.56	-0.31
	D10_22	958	8	40.37	4.78	29.5 - 69.6	2	2.48	4.56	0.34
	D15_18	960	8	40.37	3.9	30.5 - 72.6	1	1.91	-	0.45
	D18_22	959	1	52.42	3.86	48 - 60	3	1.59	5.73	0.27
	D18_22	959	3	53.58	2.9	47 - 82.7	3	1.17	5.73	-0.23
	D18_22	959	8	23.21	5.64	17 - 40.5	3	2.65	5.73	0.35
	D22_27	957	1	65.9	3.95	48.5 - 74.2	5	1.91	10.09	0.29
	D22_27	957	3	66.24	5.95	55 - 82.7	5	2.64	10.09	-0.34
	D22_27	957	7	41.08	4.77	30.5 - 48.5	5	2.12	10.09	-0.31
	D22_27	957	9	27.5	4.01	22.5 - 37	5	1.85	10.09	0.29
	D22_27	957	10	3	3.72	0 - 10.5	5	1.69	10.09	0.28
	D22_30	956	1	65.9	4.23	48.9 - 74.2	4	1.78	6.98	0.24
	D22_30	956	3	73.34	3.18	45 - 87	4	1.38	6.98	-0.21
	D22_30	956	7	38.89	3	27 - 49	4	1.22	6.98	-0.2
	D22_30	956	9	31.71	4.39	22 - 36	4	2.39	6.98	0.27
	D30_35	957	1	65.99	3.5	51 - 71.7	4	1.6	7.79	0.19
	D30_35	957	5	60.34	3.79	34.5 - 65.5	4	1.71	7.79	-0.19
	D30_35	957	7	33.43	4.23	21.5 - 46	4	2.05	7.79	-0.21
	D30_35	957	8	31.11	5.89	20.5 - 40.5	4	2.51	7.79	0.23
	IDM	944	9	37.5	4.67	22.5 - 41	1	2.29	-	0.06

Nombre de génotype (n), Groupe de liaison (LG), position de l'optimum du QTL (cM), intervalle de confiance (BCI), pourcentage de variance expliqué par un QTL (PEV, %) et effet de substitution allélique (D).



Tableau S6.4 : Moyennes et intervalles sur les deux années d'essai des coefficients de corrélation (Pearson) entre les paramètres journaliers pour l'ensemble des arbres (AMP ou ER) et les variables environnementales journalières. Les coefficients de corrélation sont calculés sur deux périodes : la semaine ou le mois. La valeur journalière est le maximum pour TC, la moyenne pour VPD, RG et REW, et la somme pour PP.

		AMP			ER		
		Moyenne (écart type)	Intervalle	%	Moyenne (écart type)	Intervalle	%
TC	Semaine	0.5 (0.17)	-0.02 , 0.85	88.7	-0.07 (0.18)	-0.53 , 0.32	78.3
TC	Mois	0.52 (0.1)	0.33 , 0.72	100	-0.04 (0.15)	-0.42 , 0.2	83.3
VPD	Semaine	0.57 (0.15)	0.21 , 0.88	93.4	-0.12 (0.18)	-0.56 , 0.3	79.2
VPD	Mois	0.61 (0.09)	0.42 , 0.84	100	-0.09 (0.16)	-0.43 , 0.17	91.7
RG	Semaine	0.56 (0.17)	-0.18 , 0.87	12.3	-0.13 (0.2)	-0.63 , 0.36	72.6
RG	Mois	0.58 (0.1)	0.39 , 0.8	4.2	-0.08 (0.15)	-0.4 , 0.18	91.7
PP	Semaine	-0.26 (0.25)	-0.82 , 0.35	97.2	0.1 (0.21)	-0.32 , 0.57	81.1
PP	Mois	-0.25 (0.12)	-0.54 , -0.02	100	0.11 (0.14)	-0.24 , 0.36	91.7
REW	Semaine	0.01 (0.29)	-0.63 , 0.77	34.9	-0.06 (0.2)	-0.71 , 0.52	77.4
REW	Mois	-0.08 (0.16)	-0.3 , 0.27	25	-0.04 (0.14)	-0.41 , 0.23	87.5

Tableau S6.5 : Moyenne, l'écart type et l'intervalle sur la période d'étude des coefficients de corrélation (Pearson) entre les paramètres journaliers (AMP ou ER) et les variables environnementales journalières par génotype. Les coefficients de corrélation sont calculés à l'échelle de la journée, de la semaine et du mois. Pour ces deux dernières échelles la moyenne (m) et le cumul (c) sont présentés.

			AMP			ER		
Variable	Période		Moyenne	Intervalle	%	Moyenne	Intervalle	%
TC	Jour	m	0.39 (0.13)	-0.04 , 0.68	99.1	0.25 (0.12)	-0.17 , 0.51	94.8
TC	Semaine	m	0.27 (0.19)	-0.27 , 0.69	64.2	0.32 (0.14)	-0.15 , 0.64	84
TC	Semaine	c	0.25 (0.19)	-0.2 , 0.72	59.9	0.22 (0.14)	-0.17 , 0.63	56.1
TC	Mois	m	0.24 (0.23)	-0.3 , 0.77	25.9	0.34 (0.15)	-0.19 , 0.69	35.8
TC	Mois	c	0.16 (0.24)	-0.4 , 0.73	16	0.25 (0.16)	-0.18 , 0.71	15.1
VPD	Jour	m	0.66 (0.09)	0.4 , 0.86	99.5	0.01 (0.08)	-0.24 , 0.22	30.7
VPD	Semaine	m	0.56 (0.14)	0.14 , 0.85	99.1	0.11 (0.12)	-0.21 , 0.43	21.2
VPD	Semaine	c	0.5 (0.14)	0.03 , 0.77	97.6	0.1 (0.12)	-0.21 , 0.42	22.6
VPD	Mois	m	0.48 (0.19)	-0.06 , 0.88	66	0.15 (0.15)	-0.23 , 0.51	4.7
VPD	Mois	c	0.44 (0.18)	-0.11 , 0.84	57.5	0.14 (0.15)	-0.24 , 0.49	3.3
RG	Jour	m	0.47 (0.13)	0.11 , 0.79	99.5	0.18 (0.09)	-0.12 , 0.42	92.9
RG	Semaine	m	0.3 (0.2)	-0.17 , 0.8	67.9	0.38 (0.14)	-0.29 , 0.72	92.5
RG	Semaine	c	0.27 (0.21)	-0.18 , 0.81	61.8	0.36 (0.13)	-0.03 , 0.66	91
RG	Mois	m	0.25 (0.26)	-0.34 , 0.87	31.1	0.41 (0.16)	-0.36 , 0.79	53.8
RG	Mois	c	0.21 (0.28)	-0.37 , 0.88	30.2	0.41 (0.14)	-0.02 , 0.74	51.4
PP	Jour	m	-0.19 (0.04)	-0.27 , -0.08	99.1	0.16 (0.06)	-0.02 , 0.31	92
PP	Semaine	m	-0.19 (0.11)	-0.42 , 0.18	52.8	0.34 (0.11)	-0.02 , 0.57	90.6
PP	Semaine	c	-0.19 (0.12)	-0.48 , 0.28	54.2	0.32 (0.11)	0.01 , 0.58	89.6
PP	Mois	m	-0.09 (0.18)	-0.47 , 0.41	1.9	0.42 (0.14)	-0.05 , 0.75	59
PP	Mois	c	-0.1 (0.19)	-0.5 , 0.52	3.8	0.41 (0.14)	0.05 , 0.78	53.3
REW	Jour	m	-0.12 (0.12)	-0.44 , 0.29	69.8	0.19 (0.11)	-0.14 , 0.57	85.8
REW	Semaine	m	-0.2 (0.17)	-0.59 , 0.34	51.9	0.25 (0.13)	-0.13 , 0.66	68.4
REW	Semaine	c	-0.16 (0.17)	-0.52 , 0.35	46.2	0.25 (0.13)	-0.12 , 0.66	66
REW	Mois	m	-0.21 (0.21)	-0.75 , 0.44	19.8	0.25 (0.15)	-0.15 , 0.75	14.6
REW	Mois	c	-0.17 (0.22)	-0.65 , 0.43	14.6	0.27 (0.15)	-0.13 , 0.75	17.5



Tableau S6.6 : QTL détectés pour les paramètres AMP et ER et leurs corrélations avec les variables environnementales (TC, RG, VPD, PP, REW). Les paramètres à la journée (j) ou cumulés sur la semaine (sc) ou le mois (mc) sont moyennés sur toute la période d'essai. Les corrélations sont celles sur les données journalière (J) ou les moyennes (Sm) ou les cumuls (Sc) par semaine

Parent	Caractère	n	LG	Position (cM)	LOD	BCI	Modèle QTL	PEV	PEV tot	D
<i>E.g</i>	AMPj	210	8	5.32	5.51	1.3 - 14.5	1	7.68		11
	AMPsc	210	8	4.73	5.05	0.5 - 14.5	1	6.55		81
	AMPmc	210	8	4.73	5.06	0.5 - 14.5	1	6.55		351
	J.TC.a	210	2	55	3.74	44.5 - 78.7	1	6.92		-0.07
	Sm.TC.a	210	2	76.51	2.89	46 - 87.5	1	5.25		-0.09
	J.VPD.a	210	2	54.1	4.73	46 - 64	1	9.08		-0.06
	Sm.VPD.a	210	2	54.6	3.84	46 - 64.5	1	8.11		-0.08
	Sc.VPD.a	210	2	55.28	4.34	46 - 66.5	1	6.88		-0.07
	J.PP.a	210	4	56.78	2.95	39 - 61.9	2	3.66	10.41	-0.01
	J.PP.a	210	8	27.73	3.48	13.5 - 35.5	2	7.6	10.41	-0.02
<i>E.u</i>	J.PP.e	210	9	55.9	3.1	33.5 - 55.9	1	4.85		0.03
	ERt	210	5	8.19	3.13	5.5 - 20.1	1	6.09		5606
	J.TC.a	210	3	5.82	3.4	0 - 13.5	1	5.61		-0.06
	Sc.TC.a	210	3	5.82	4.42	0 - 11.6	1	6.36		-0.1
	Sc.VPD.a	210	3	5.82	3.37	0 - 12.7	1	6.01		-0.07
	J.RG.a	210	3	5.92	3.46	0 - 16.5	1	5.86		-0.06
	Sm.RG.a	210	6	41.36	3.71	0.3 - 44.8	1	5.9		0.1
	Sc.RG.a	210	3	11.64	3.03	0 - 17	2	4.96	11.02	-0.09
	Sc.RG.a	210	6	41.36	3.71	5.5 - 44.8	2	6.22	11.02	0.11
	Sc.PP.a	210	2	76	2.91	68.5 - 98	1	4.01		-0.05
	J.REW.a	210	2	86.52	3.5	69.3 - 103.2	1	6.62		-0.06
	Sm.REW.a	210	2	86.52	3.16	37.3 - 103.2	2	6.4	11.4	-0.09
	Sm.REW.a	210	4	55.03	2.96	21 - 56.5	2	5.16	11.4	0.08
	Sc.REW.a	210	2	74.53	4.51	71.5 - 90.5	1	6.13		-0.09
	J.TC.e	210	2	86.52	2.94	61.9 - 92	1	4.5		0.05
	Sc.TC.e	210	2	86.52	3.12	67.4 - 92.5	1	4.24		0.06
	J.VPD.e	210	4	6.61	4.2	0 - 17.1	1	3.86		0.03
	Sm.VPD.e	210	4	6.61	3.98	0 - 19.6	1	4.06		0.05
	J.RG.e	210	4	3.06	5.15	0 - 13	1	6.5		0.05
	Sm.RG.e	210	4	5.13	3.83	0 - 18.8	1	6.26		0.07
	J.PP.e	210	10	30.57	2.86	1.2 - 53	1	3.53		-0.02
	J.REW.e	210	2	61.91	4.28	56.5 - 87.3	1	6.95		0.06
	Sm.REW.e	210	2	61.91	3.51	54.8 - 87.5	1	6.57		0.07
	Sc.REW.e	210	2	61.91	3.64	57 - 87.5	1	6.18		0.06

Nombre de génotype (n), Groupe de liaison (LG), position de l'optimum du QTL (cM), intervalle de confiance (BCI), pourcentage de variance expliqué par un QTL (PEV, %) et effet de substitution allélique (D).



Tableau S6.7 : QTL détectés pour les paramètres AMP et ER sur deux semaines contrastées : une de saison des pluies (S1) et une de saison sèche (S2). Le chiffre après le paramètre indique le jour de la semaine, lorsqu'il n'y en pas il s'agit de la moyenne sur la semaine.

Parent	Semaine	Caractère	n	LG	Position (cM)	LOD	BCI	Modèle QTL	PEV	PEV tot	D
<i>E.g</i>	S2	2.AMP1	210	8	25.76	3.88	0 - 31	1	6.78		27.44
		2.AMP2	210	8	25.76	3.94	4 - 32	1	7.52		32.69
		2.AMP3	210	8	25.76	3.19	1.3 - 34	1	5.17		19.92
		2.AMP6	210	8	25.76	3.73	2 - 33.5	1	7.78		25.43
		2.AMP7	210	8	25.76	3.01	3.5 - 46.3	1	6.02		15.64
		2.AMP	210	8	25.76	3.53	2 - 32	1	7.12		23.19
		2.ER3	210	10	48.5	2.89	31.5 - 50.5	1	7.23		-10.74
		2.ER5	210	8	32.38	4.8	29 - 44.9	1	7.26		10.77
<i>E.u</i>	S1	1.AMP1	210	3	10.85	4.68	4.5 - 14	2	9.43	13.02	-24.43
		1.AMP1	210	9	17.98	3.14	1.5 - 22.5	2	4.61	13.02	-16.99
		1.AMP2	210	3	12	5.41	5 - 14	1	7.85		-30.8
		1.AMP3	210	3	12.73	3.94	3.5 - 26	1	5.98		-29.43
		1.AMP4	210	3	10.85	5.69	4.7 - 13.8	1	8.91		-24.8
		1.AMP5	210	3	10.5	5.31	4.7 - 13.8	1	8.26		-29.82
		1.AMP6	210	3	10.85	3.79	0.5 - 14	2	6.73	12.83	-31.28
		1.AMP6	210	9	12.58	3.37	1 - 19	2	6.7	12.83	-31.09
		1.AMP7	210	3	10.85	3.11	0 - 14	1	7.63		-35.85
		1.AMP	210	3	10.85	4.99	4.5 - 14	1	8.61		-29.02
		1.ER3	210	6	65.5	3.22	50 - 81.5	1	5.07		-15.13
	S2	2.AMP1	210	3	45.02	4.57	30.5 - 59.3	2	7.75	12.94	-29.28
		2.AMP1	210	8	104.14	2.93	88.5 - 104.1	2	5.35	12.94	-24.37
		2.AMP2	210	3	45.02	4.79	30.5 - 59	1	7.38		-32.34
		2.AMP3	210	3	45.02	4.3	33.7 - 60	1	7.1		-23.28
		2.AMP4	210	3	45.02	5.33	35 - 59	1	7.67		-26.45
		2.AMP5	210	3	45.02	4.18	33.7 - 80.6	1	7.43		-19.71
		2.AMP6	210	3	43.26	5.05	35 - 59.5	1	8.02		-25.74
		2.AMP7	210	3	72.85	3	35.2 - 86	1	4.14		-12.95
		2.AMP	210	3	45.02	5.1	33.7 - 59.5	1	8.35		-25.06

Nombre de génotype (n), Groupe de liaison (LG), position de l'optimum du QTL (cM), intervalle de confiance (BCI), pourcentage de variance expliqué par un QTL (PEV, %) et effet de substitution allélique (D).





## Références

1. Kozłowski TT, Winget CH: **Diurnal and Seasonal Variation in Radii of Tree Stems.** *Ecology* 1964, **45**(1):149-155.
2. Drew DM, Downes GM: **The use of precision dendrometers in research on daily stem size and wood property variation: A review.** *Dendrochronologia* 2009, **27**(2):159-172.
3. Herzog K, Häslér R, Thum R: **Diurnal changes in the radius of a subalpine Norway spruce stem: their relation to the sap flow and their use to estimate transpiration.** *Trees* 1995, **10**(2):94-101.
4. Offenthaler I, Hietz P, Richter H: **Wood diameter indicates diurnal and long-term patterns of xylem water potential in Norway spruce.** *Trees - Structure and Function* 2001, **15**(4):215-221.
5. Downes G, Beadle C, Worledge D: **Daily stem growth patterns in irrigated Eucalyptus globulus and E. nitens in relation to climate.** *Trees* 1999, **14**(2):102-111.
6. Deslauriers A, Morin H, Urbinati C, Carrer M: **Daily weather response of balsam fir (Abies balsamea (L.) Mill.) stem radius increment from dendrometer analysis in the boreal forests of Québec (Canada).** *Trees* 2003, **17**(6):477-484.
7. Bouriaud O, Leban J-M, Bert D, Deleuze C: **Intra-annual variations in climate influence growth and wood density of Norway spruce.** *Tree Physiology* 2005, **25**(6):651-660.
8. Zweifel R, Item H, Häslér R: **Stem radius changes and their relation to stored water in stems of young Norway spruce trees.** *Trees* 2000, **15**(1):50-57.
9. Zweifel R, Item H, Häslér R: **Link between diurnal stem radius changes and tree water relations.** *Tree Physiology* 2001, **21**(12-13):869-877.
10. Cochard H, Forestier S, Améglio T: **A new validation of the Scholander pressure chamber technique based on stem diameter variations.** *Journal of Experimental Botany* 2001, **52**(359):1361-1365.
11. Deslauriers A, Anfodillo T, Rossi S, Carraro V: **Using simple causal modeling to understand how water and temperature affect daily stem radial variation in trees.** *Tree Physiology* 2007, **27**(8):1125-1136.
12. King G, Fonti P, Nievergelt D, Büntgen U, Frank D: **Climatic drivers of hourly to yearly tree radius variations along a 6°C natural warming gradient.** *Agricultural and Forest Meteorology* 2013, **168**(0):36-46.
13. Deslauriers A, Rossi S, Anfodillo T: **Dendrometer and intra-annual tree growth: What kind of information can be inferred?** *Dendrochronologia* 2007, **25**(2):113-124.
14. Mareschal L, Nzila JDD, Turpault MP, Thongo M'Bou A, Mazoumbou JC, Bouillet JP, Ranger J, Laclau JP: **Mineralogical and physico-chemical properties of Ferralic Arenosols derived from unconsolidated Plio-Pleistocene deposits in the coastal plains of Congo.** *Geoderma* 2011, **162**(1-2):159-170.
15. IUSS Working Group: **World reference base for soil resources 2006.** In: *World Soil Resources Reports*. ROME: FAO; 2006.
16. Bouillet J-P, Laclau J-P, Arnaud M, M'Bou AT, Saint-André L, Jourdan C: **Changes with age in the spatial distribution of roots of Eucalyptus clone in Congo: Impact on water and nutrient uptake.** *Forest Ecology and Management* 2002, **171**(1-2):43-57.
17. Villar E: **Plasticité phénotypique et moléculaire de deux clones d'eucalyptus sous contrainte hydrique au champ.** *Ph.D.* Bordeaux: Université de Bordeaux I; 2011.
18. R Core Team: **R: A Language and Environment for Statistical Computing.** In: R Foundation for Statistical Computing, Vienna, Austria: URL <http://www.R-project.org/>; 2013.
19. Wickham H: **ggplot2: Elegant Graphics for Data Analysis.** Springer Publishing Company, Incorporated; 2009.
20. Grattapaglia D, Sederoff R: **Genetic Linkage Maps of Eucalyptus grandis and Eucalyptus urophylla Using a Pseudo-Testcross: Mapping Strategy and RAPD Markers.** *Genetics* 1994, **137**(4):1121-1137.
21. Margarido GRA, Souza AP, Garcia AAF: **OneMap: software for genetic mapping in outcrossing species.** *Hereditas* 2007, **144**(3):78-79.
22. Os H, Stam P, Visser RF, Eck H: **RECORD: a novel method for ordering loci on a genetic linkage map.** *Theor Appl Genet* 2005, **112**(1):30-40.



23. Mollinari M, Margarido GRA, Vencovsky R, Garcia AAF: **Evaluation of algorithms used to order markers on genetic maps.** *Heredity* 2009, **103**(6):494-502.
24. Broman KW, Wu H, Sen Ś, Churchill GA: **R/qtl: QTL mapping in experimental crosses.** *Bioinformatics* 2003, **19**(7):889-890.
25. Haley CS, Knott SA: **A simple regression method for mapping quantitative trait loci in line crosses using flanking markers.** *Heredity* 1992, **69**(4):315-324.
26. Churchill GA, Doerge RW: **Empirical Threshold Values for Quantitative Trait Mapping.** *Genetics* 1994, **138**(3):963-971.
27. Voorrips RE: **MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs.** *Journal of Heredity* 2002, **93**(1):77-78.
28. Tardif J, Flannigan M, Bergeron Y: **An Analysis of the Daily Radial Activity of 7 Boreal Tree Species, Northwestern Quebec.** *Environ Monit Assess* 2001, **67**(1-2):141-160.
29. Devine WD, Harrington CA: **Factors affecting diurnal stem contraction in young Douglas-fir.** *Agricultural and Forest Meteorology* 2011, **151**(3):414-419.
30. Daudet F-A, Améglio T, Cochard H, Archilla O, Lacointe A: **Experimental analysis of the role of water and carbon in tree stem diameter variations.** *Journal of Experimental Botany* 2005, **56**(409):135-144.
31. Drew D, Downes G, Grzeskowiak V, Naidoo T: **Differences in daily stem size variation and growth in two hybrid eucalypt clones.** *Trees - Structure and Function* 2009, **23**(3):585-595.
32. Zweifel R, Zimmermann L, Newbery DM: **Modeling tree water deficit from microclimate: an approach to quantifying drought stress.** *Tree Physiology* 2005, **25**(2):147-156.
33. Zweifel R, Zimmermann L, Zeugin F, Newbery DM: **Intra-annual radial growth and water relations of trees: implications towards a growth mechanism.** *Journal of Experimental Botany* 2006, **57**(6):1445-1459.
34. Wimmer R, Downes GM, Evans R: **Temporal variation of microfibril angle in Eucalyptus nitens grown in different irrigation regimes.** *Tree Physiology* 2002, **22**(7):449-457.
35. Drew D, Downes G, Evans R: **Short-term growth responses and associated wood density fluctuations in variously irrigated Eucalyptus globulus.** *Trees - Structure and Function* 2011, **25**(2):153-161.
36. Beavis WD: **QTL analyses: power, precision, and accuracy.** *Molecular dissection of complex traits* 1998, **1998**:145-162.
37. Xu S: **Theoretical Basis of the Beavis Effect.** *Genetics* 2003, **165**(4):2259-2268.
38. Welcker C, Sadok W, Dignat G, Renault M, Salvi S, Charcosset A, Tardieu F: **A Common Genetic Determinism for Sensitivities to Soil Water Deficit and Evaporative Demand: Meta-Analysis of Quantitative Trait Loci and Introgression Lines of Maize.** *Plant Physiol* 2011, **157**(2):718-729.
39. Vigneron P: **Création et amélioration de variétés hybrides d'Eucalyptus au Congo.** *Français* 1992, **234**:29-42.
40. Gion J-M, Carouche A, Deweer S, Bedon F, Pichavant F, Charpentier J-P, Bailleres H, Rozenberg P, Carocha V, Ognouabi N *et al*: **Comprehensive genetic dissection of wood properties in a widely-grown tropical tree: Eucalyptus.** *BMC Genomics* 2011, **12**(1):301.
41. Wimmer R, Downes GM, Evans R: **High-resolution analysis of radial growth and wood density in Eucalyptus nitens, grown under different irrigation regimes.** *Ann For Sci* 2002, **59**(5-6):519-524.
42. Zweifel R, Häsler R: **Dynamics of water storage in mature subalpine Picea abies: temporal and spatial patterns of change in stem radius.** *Tree Physiology* 2001, **21**(9):561-569.





## **7 Conclusion et perspectives**



L'objectif de ces travaux de thèse était de caractériser le déterminisme génétique et environnemental de la croissance de l'eucalyptus, grâce à l'analyse de l'architecture génétique de la dynamique de croissance et de la composition isotopique du carbone ( $\delta^{13}\text{C}$ ). Les analyses ont été menées sur un croisement hybride interspécifique *E. urophylla* x *E. grandis*, croisement de référence pour les travaux de cartographie génétique et d'analyse QTL sur eucalyptus au CIRAD. Ce croisement a été étudié à partir de quatre dispositifs expérimentaux présents à Pointe-Noire (République du Congo). Notre approche combine i/ l'utilisation de technologies de génotypage haut débit sur un grand nombre d'hybrides, ii/ l'ancrage de carte génétiques sur un génome de référence, iii/ des données de croissance classiquement utilisées au sein du programme d'amélioration, iv/ ainsi que des données de phénotypage en continu de la micro-variation radiale (inédite à cette échelle chez les arbres forestiers). Cette approche nous a permis d'apporter des résultats originaux sur l'architecture génétique de la croissance et les réponses de l'arbre face aux contraintes abiotiques. Ce dernier chapitre dresse une synthèse des principaux résultats issus de ces travaux ainsi que certaines limites, et présente les perspectives qui y sont associées.

## 7.1 Synthèse des résultats obtenus

### 7.1.1 Les apports des cartes génétiques haute résolution

Les cartes génétiques sont la base de l'analyse QTL et sont donc au cœur des travaux présentés dans cette thèse. Les travaux sur l'architecture des caractères complexes chez les arbres forestiers ont longtemps été menés à partir de cartes génétiques à faible densité de marqueurs [1-9]. Grâce à la démocratisation des technologies de génotypage haut débit, ils sont aujourd'hui conduits à partir de cartes génétiques à haute densité [10-12]. Cependant, les types de marqueurs utilisés, ainsi que les faibles effectifs utilisés au sein des descendance de cartographie, sont deux facteurs qui peuvent limiter la précision de ce type d'analyse. En effet, la couverture du génome n'est pas forcément optimale (génotypage de marqueurs sans connaissance *a priori* de leur localisation sur le génome) et la précision de l'estimation des liaisons entre marqueurs peut être faible. Les cartes génétiques "haute résolution" développées dans le cadre de cette thèse ont donc eu pour objectif de lever ces deux verrous.

L'approche de cartographie génétique proposée dans le chapitre 4 a été possible grâce au reséquençage des deux parents du croisement afin de mettre en évidence l'ensemble du polymorphisme nucléotidique porté par ceux-ci. Grâce à cette information, une sélection de



six milles SNP répartis le long du génome de référence a été effectuée. Le génotypage a révélé 4367 SNP polymorphes dans le croisement, offrant ainsi une couverture d'environ 98% du génome (les onze principaux "scaffolds" correspondant aux chromosomes), avec en moyenne un SNP tous les 140 kb. L'utilisation d'un large effectif nous a permis de capturer un nombre important d'évènements de recombinaison, et donc d'estimer précisément les distances génétiques. Ce changement d'échelle a permis de mettre en évidence que les tailles des cartes génétiques rapportées dans les études récentes chez l'*Eucalyptus* [10, 13, 14] sont probablement surestimées d'environ 10%. Les cartes obtenues indépendamment pour les deux parents montrent une colinéarité importante entre les génomes d'*E. grandis* et d'*E. urophylla*, comme suggéré par Kullán *et al.* [15]. Par ailleurs, grâce à leur ancrage sur le génome de référence BRASUZ1, elles ont permis de comparer l'ordre des marqueurs entre carte physique et cartes génétiques. Des différences entre la localisation génétique et la position physique sur le génome de référence BRASUZ1 ont été mises en évidence. Elles sont au nombre de 56 et représentent 15% de la taille du génome. Ce sont majoritairement des inversions au sein des "scaffolds", mais des inversions entre "scaffolds" ont également été trouvées. La validation croisée grâce aux deux cartes parentales nous permet de penser que ces différences sont certainement dues à des erreurs d'assemblage dans la première version. En effet l'assemblage du génome de référence est basé sur des cartes consensus avec des marqueurs DArT [14, 15]. Il est possible, compte tenu de l'effectif utilisé pour construire ces cartes (367 et 180 individus), que certains marqueurs aient été mal positionnés. Ce type d'inversion est fréquemment rapporté lors de la production de génome chez différentes espèces comme *Zea mays* [16], *Solanum tuberosum* [17], *Brassica rapa* [18] ou *Citrullus lanatus* [19]. Nos résultats vont donc permettre d'améliorer l'assemblage de la première version du génome de référence. La qualité de l'assemblage est un point crucial pour de nombreuses analyses basées sur le génome.

### 7.1.2 Architecture génétique de la dynamique de croissance

L'architecture génétique de la croissance chez les arbres forestiers a été largement étudiée au cours des deux dernières décennies [20, 21]. La plupart des études rapportent une instabilité des QTL de croissance dans le temps [1, 3, 8, 22], mais aucune chez l'*Eucalyptus* n'a étudié l'évolution de cette instabilité sur l'ensemble de la durée de rotation. Les travaux décrits dans le chapitre 3 (dispositif P93, P97 et P98), combinent plusieurs approches d'analyse des caractères de croissance (hauteur et circonférence). La caractérisation de



l'environnement, combinée à celle de la dynamique de croissance sur la durée de rotation, nous a permis de valider une des hypothèses sur l'intérêt de la prise en compte de la dynamique (c.à.d. des trajectoires), afin de décomposer les réponses aux variations environnementales. Bien qu'ayant une faible résolution temporelle, l'indice climatique utilisé (indice de "de Martonne" [23]) a permis de montrer le lien entre la disponibilité en eau et la croissance au jeune âge (avant deux ans). L'alternance entre des périodes de contrainte hydrique marquée et des périodes de forte disponibilité en eau semble structurer le classement des génotypes, en favorisant ceux qui répondent le mieux à cette alternance. On peut rapprocher ce résultat de ceux obtenus sur deux clones d'*Eucalyptus* ssp. en République du Congo, où le clone le plus productif est également celui qui répond (reprise de croissance) le plus vite au changement de saison [24].

L'analyse QTL à partir des accroissements a révélé une architecture génétique plus complexe (nombre de régions génomique plus important) que celle observée avec les caractères de croissance cumulée, montrant ainsi une variété des réponses sous-jacentes à la croissance totale. Cependant, un déterminisme génétique commun a été mis en évidence entre la réponse au jeune âge (accroissement) et la croissance cumulée à cinq ans. Ce résultat, ainsi que la forte interaction QTL×E mise en évidence entre dispositifs nous a permis de conclure que la plasticité de la croissance au stade juvénile détermine en partie l'architecture génétique de croissance à l'âge de la rotation.

Les résultats sur un effectif plus large de la même famille (chapitre 6, dispositif P10) renforcent cette conclusion. En effet, l'architecture génétique de la croissance totale à trois ans a été reliée à celle des accroissements en début de plantation (avant deux ans). Ce changement d'échelle au niveau du nombre de descendants a également permis de préciser les effets et le nombre des QTL impliqués. Le nombre de QTL détecté a été multiplié par deux et les effets moyens divisés par 3 par rapport à la première étude (chapitre 3) réalisée sur environ 200 individus. Cependant, ce changement d'échelle n'a pas d'incidence sur les intervalles de confiance des QTL. Ceci peut s'expliquer par l'héritabilité faible à moyenne des caractères de croissance [25, 26]. La prise en compte de l'environnement, grâce à des répétitions clonales pourrait avoir une incidence sur les intervalles de confiance plus important que la seule augmentation du nombre de recombinants observables dans la population.

Cette première partie sur l'architecture génétique de la croissance nous a permis de mettre en évidence une variabilité génétique de la plasticité phénotypique (G×E) en réponse à





l'alternance des saisons. Par la suite, on s'est attaché à caractériser plus finement cette part génétique de la plasticité phénotypique ainsi que ses déterminants abiotiques.

### 7.1.3 Architecture génétique de la composition isotopique du carbone

L'intégration des réponses à l'environnement a également été abordée en utilisant un caractère classiquement utilisé pour étudier la croissance en situation de contrainte hydrique : la composition isotopique du carbone ( $\delta^{13}\text{C}$ ) [27, 28]. Différentes échelles d'intégration (saison et durée de rotation) du  $\delta^{13}\text{C}$  ont été étudiées sur trois dispositifs. Cela a permis de caractériser pour la première fois, la variabilité du  $\delta^{13}\text{C}$  au niveau d'une famille de plein-frères d'eucalyptus ainsi que son architecture génétique (chapitre 5).

Une gamme de variation de 4‰ a été observée dans la famille. Cette gamme est relativement proche de ce que l'on peut rencontrer à l'échelle du genre *Eucalyptus*. En effet, Schulze *et al.* [29] ont rapporté une gamme de variation de 5,2 ‰ entre 65 espèces d'eucalyptus, prélevées le long d'un gradient de pluviométrie. La variabilité pour le  $\delta^{13}\text{C}$  n'a pas été mise en lien avec celle pour la croissance, et ce pour les différentes échelles d'intégrations. En effet, les corrélations entre les deux caractères sont majoritairement non significatives ou inférieures à 0.3. Ce manque de corrélation a été rapporté chez d'autres espèces forestières [30-33] mais aucune tendance claire ne se dégage de la littérature [34-39]. Bien que le manque de corrélation soit confirmé par les trois dispositifs, il conviendrait de valider ce résultat par l'étude d'un pool génétique plus large, comme celui présent au sein du schéma de sélection récurrente-réciproque basé sur le croisement interspécifique *E. urophylla* et *E. grandis*.

La variabilité phénotypique importante au sein de la famille a été associée à des régions génomiques stables entre dispositifs, notamment entre P93 et P97. Cette stabilité entre dispositifs peut s'expliquer par la durée d'intégration associée aux mesures de  $\delta^{13}\text{C}$  sur holocellulose extraite de bois, qui minimise les effets environnementaux. À l'inverse, aucune colocalisation n'a été trouvée entre les QTL de  $\delta^{13}\text{C}$  de saison sèche et ceux de saison des pluies, suggérant un effet marqué du déficit hydrique sur le déterminisme génétique du  $\delta^{13}\text{C}$ , à l'instar des résultats sur *Salix* [27] et *Vitis* [28]. Néanmoins, l'effet du changement progressif de la structure des feuilles (hétéroblastie foliaire) marqué au jeune âge chez l'eucalyptus (période de prélèvement) peut également jouer un rôle dans ce changement de classement. Ce résultat souligne que le  $\delta^{13}\text{C}$ , bien qu'étant un caractère intégrateur de plusieurs types de réponses (à l'image de la croissance), a une architecture génétique qui présente une forte



interaction avec l'environnement. Afin d'apporter un complément de réponse à l'effet de la saison sur le déterminisme génétique du  $\delta^{13}\text{C}$ , nous disposons d'un prélèvement supplémentaire de feuilles du dispositif P10, effectué lors de la deuxième saison sèche à 22 mois. Ce prélèvement n'était initialement pas prévu et il n'a pas été possible de l'intégrer dans la thèse compte tenu du temps imparti. L'analyse de ce prélèvement pourrait permettre une meilleure compréhension de l'effet de la saison sèche sur l'architecture génétique du  $\delta^{13}\text{C}$ , et confirmer ou non le lien possible avec le changement de structure foliaire.

Par ailleurs, une région majeure impliquée dans le déterminisme génétique du  $\delta^{13}\text{C}$  a également été mise en évidence sur le LG4. L'analyse d'enrichissement en termes de 'Gene Ontology' (GO) dans cette région a mis en évidence des gènes associés à l'ouverture/fermeture des stomates [40, 41]. Ces résultats sur l'architecture génétique du  $\delta^{13}\text{C}$  et l'analyse sans *a priori* des gènes sous-jacents à des QTL sont les premiers chez l'*Eucalyptus*. Ils sont très encourageants et montrent la puissance des cartes génétiques haute résolution combinées au génome de référence pour la caractérisation des régions génomiques impliquées dans le déterminisme génétique des caractères d'intérêts.

Enfin, les résultats du chapitre 5 montrent une architecture génétique distincte entre le  $\delta^{13}\text{C}$  et la croissance. Ceci nous a amené, dans un premier temps, à conclure que les facteurs génétiques sous-jacents à la variabilité du  $\delta^{13}\text{C}$  (et par extension l'efficacité d'utilisation de l'eau) et de la croissance étaient découplés. Cependant les analyses du chapitre 6 sur l'architecture génétique de la croissance avec un effectif plus large (960 individus) ont permis de détecter des QTL d'accroissement en diamètre (non détecté avec 200 individus) qui, colocalisent avec un QTL de  $\delta^{13}\text{C}$  sur le LG4 (*E. urophylla*) et sur le LG8 (*E. grandis*). La précision accrue grâce à l'augmentation du nombre de génotypes peut expliquer cette colocalisation. Ce dernier résultat pourrait nous amener à infirmer l'hypothèse formulée sur un déterminisme génétique différent entre croissance et  $\delta^{13}\text{C}$ . Toutefois, nous n'avons pas observé de corrélation phénotypique entre ces caractères de croissance et le  $\delta^{13}\text{C}$  et de plus, les LOD maximum de ces QTL ne sont pas communs. Cela suggérerait donc des gènes liés plutôt qu'un gène à effet pléiotropique, contrairement à l'hypothèse formulée par Rönnerberg-Wästljung et al. [27] sur des cartes à faibles densités.



#### 7.1.4 Architecture de la plasticité des micro-variations du rayon aux variations environnementales

Parallèlement aux mesures ponctuelles de croissance sur le dispositif P10 (chapitre 6), nous avons cherché à caractériser plus finement la réponse, i.e. la plasticité phénotypique, de l'arbre à l'environnement en début de plantation. Pour ce faire, un dispositif de suivi en continu des variations du rayon du tronc a donc été mis en place en au début de cette thèse et a constitué un défi technique (chapitre 2). Grâce à ce suivi continu de la croissance, nous avons pu pousser la logique de décomposition des réponses à l'environnement jusqu'à l'échelle journalière.

Les variations journalières du rayon du tronc sont connues pour intégrer différents processus dont les deux principaux sont : i/ le rétrécissement puis le gonflement des tissus périphériques (principalement le cambium et le phloème), liés au déstockage et stockage d'eau du tronc, et ii/ la croissance (mise en place de tissu au niveau du cambium) [42-51]. Deux paramètres ont été extraits de ces variations journalières du rayon, l'un associé au fonctionnement hydrique de l'arbre : l'amplitude journalière et l'autre associé à la croissance : l'évolution nette du rayon. Ces deux caractères non corrélés montrent des différences de réponse à l'environnement. En effet, l'amplitude s'est révélée être, en moyenne sur l'ensemble des arbres, positivement corrélée aux variables "atmosphériques" (déficit de pression de vapeur d'eau, le rayonnement et la température) à l'échelle de la journée. Ce résultat, qui était attendu du fait du lien entre la transpiration et l'amplitude journalière, a largement été rapporté dans la littérature [43, 52]. À l'inverse l'évolution nette du rayon a montré de faibles corrélations avec les variables environnementales à l'échelle de la journée. Avec l'augmentation du niveau d'intégration (semaine), des corrélations plus fortes avec la pluviométrie et le rendement global notamment, ont été mises en évidence pour ce caractère. L'effet du climat à plus long terme sur l'évolution nette du rayon a été décrit pour différentes espèces [47, 53, 54]. Cette effet à plus long terme des précipitations sur l'évolution nette du rayon est à relier avec les corrélations mises en évidence entre l'indice de "de Martonne" et les accroissements. Une variabilité importante entre génotypes a été associée à ces corrélations suggérant un contrôle génétique des réponses à l'environnement. Cette étude est la première à présenter une analyse génétique sur ce type de caractère, et sur un nombre d'individus aussi important. En effet, les approches écophysiologiques caractérisant ce signal ont été réalisées au maximum sur une soixantaine d'individus [45, 52, 54-56].



L'analyse QTL sur la plasticité de l'amplitude et de l'évolution nette du rayon, a montré un déterminisme génétique très différent entre ces deux caractères. Ce résultat souligne l'intérêt de décomposer la micro-variation du rayon pour l'étude de son déterminisme génétique. De plus, une architecture génétique différente entre la réponse à la demande évaporatoire et la réponse à la disponibilité en eau a été mise en évidence. Ce type de résultat étant rapporté pour la première fois chez une espèce forestière, il ouvre des perspectives pour une meilleure compréhension des déterminants génétiques du fonctionnement hydrique de l'arbre.

Nos résultats sur le contrôle génétique de la plasticité aux contraintes environnementales ouvrent un champ d'investigation important dans la caractérisation fonctionnelle de l'interaction GxE. En effet, l'approche d'enrichissement en GO suivie dans le cadre du  $\delta^{13}\text{C}$  pourrait être facilement étendue aux QTL de réponses issues du signal de micro-variations. Cette caractérisation des gènes candidats sous-jacents aux QTL pourraient permettre de vérifier des hypothèses, quant aux différents types de comportements associés aux différentes contraintes, en mettant en lumière les termes GO associés aux gènes sous-jacent aux QTL.

#### 7.1.5 Limites associées à ces travaux

Le choix a été fait dans cette thèse de se concentrer sur une seule famille, afin d'avoir la description la plus complète possible de la trajectoire de croissance, et ce à différentes échelles de temps. La contrepartie d'une telle approche, est la diversité génétique restreinte explorée. Plusieurs études sur *Eucalyptus* [1, 12] rapportent une instabilité de l'architecture génétique de la croissance dans différents fonds génétiques, et la généralisation de nos résultats est donc difficilement envisageable. Il serait intéressant de valider certains QTL, comme celui sur le LG4 pour le  $\delta^{13}\text{C}$  chez d'autres espèces d'*Eucalyptus*. Par ailleurs, les génotypes n'ont pas été répétés entre dispositifs QTL, ce qui ne permet pas de tester de manière formelle l'interaction GxE. Cette répétition entre dispositifs aurait permis d'analyser l'architecture génétique de l'interaction GxE, notamment au jeune âge, grâce par exemple à l'utilisation de normes de réaction comme caractère quantitatif [57]. Malgré ces limites, les travaux présentés ouvrent un grand nombre de perspectives

## 7.2 Perspectives

Les travaux présentés dans cette thèse apportent un éclairage nouveau sur les déterminants génétiques et environnementaux de la croissance chez l'*Eucalyptus* en conditions tropicales humides. Si plusieurs perspectives de recherches se dégagent de ces travaux, on insistera ici sur celles qui nous semblent possible de relever à court terme.

### 7.2.1 L'étude de la recombinaison

Au-delà de l'étude de l'architecture génétique des caractères quantitatifs, les cartes génétiques haute résolution sont un support idéal pour étudier la recombinaison le long du génome de l'*Eucalyptus* (cM/Mb). En effet, la précision d'estimation du taux de recombinaison entre marqueurs, combinée à une densité de marqueurs d'un SNP tous les 300 kb environ, est suffisantes pour caractériser les "hotspots" et les "coldspots" de recombinaison le long du génome d'*E. grandis* (Figure 7.1). Il n'existe pour l'instant aucune étude publiée chez l'*Eucalyptus* sur cette thématique, mais il existe une littérature importante pour d'autres espèces [58-62]. La recombinaison étant un processus central en amélioration des plantes (elle est à l'origine de nouvelles combinaisons d'allèles sur les chromosomes), une meilleure connaissance de la distribution des crossing-over pourrait donc permettre une efficience accrue de la sélection [63]. Ces cartes permettront donc de donner un premier aperçu de l'évolution de la recombinaison chez un arbre forestier d'intérêt économique comme l'eucalyptus. L'hétérogénéité du taux recombinaison entre parents, ainsi qu'au niveau inter et intra chromosomique, pourra être analysée en lien avec des caractéristiques structurales du génome, comme par exemple le pourcentage en guanine et cytosine (GC).

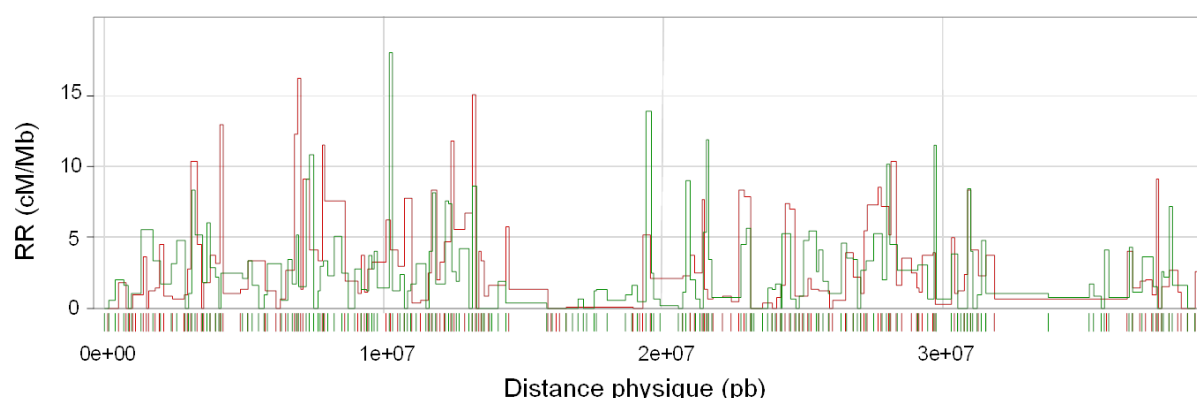


Figure 7.1 : Evolution comparée du taux de recombinaison (RR) entre *E. urophylla* (rouge) et *E. grandis* (en vert) le long du scaffold 10 du génome de BRASUZ1. Cette estimation a été obtenue à partir des cartes génétiques haute résolution présentées dans le chapitre 4.



### 7.2.2 D'autres approches d'analyse des variations journalières du rayon

Les données générées grâce au dispositif de suivi en continu des micro-variations du rayon sont abondantes (environ 250k données par arbre) et très riches. L'approche proposée dans le chapitre 6 n'est pas exhaustive, quant à la caractérisation du signal journalier de micro-variations. Plusieurs niveaux d'analyse pourraient être développés pour approfondir la caractérisation de la variabilité phénotypique et de l'architecture génétique du signal.

Au niveau du signal, d'autres paramètres comme la durée des différentes phases du cycle circadien [45, 46] ou bien une pente de contraction ou de reprise, pourraient être étudiés avec la même démarche que dans le chapitre 6. Ces paramètres supplémentaires peuvent permettre une meilleure compréhension du fonctionnement physiologique de l'arbre. Au niveau des relations entre variables environnementales et paramètres, plusieurs analyses complémentaires peuvent être faites. En effet, il serait intéressant d'analyser les corrélations ainsi que la modélisation du phénotype par saison, et non pas sur l'ensemble de la période. Cela permettrait de préciser les liens entre paramètres et variables environnementales. La modélisation du phénotype peut elle aussi être abordée sous l'angle d'un modèle mixte, incluant l'ensemble des génotypes comme effet aléatoire, ainsi qu'une pente aléatoire pour chaque variable environnementale. Enfin, au niveau de la détection QTL, la totalité des paramètres journaliers peuvent être analysés afin d'obtenir un profil de LOD score, pour un QTL donné sur l'ensemble de la période d'étude. Cette méthode permettrait d'avoir une vision en "temps réel" du déterminisme génétique de la réponse journalière à l'environnement. Par ailleurs, l'utilisation de modèles de détection de QTL de type "functional mapping" [57, 64-66], peuvent être envisagés pour étudier le déterminisme génétique du cycle circadien dans son ensemble. L'association de l'ensemble de ces méthodes permettrait d'obtenir une vision sans doute plus détaillée de l'architecture génétique sous-jacente au cycle circadien qui a été présenté dans cette thèse.

### 7.2.3 Le lien entre micro-variations et propriétés du bois

Dans cette étude, les micro-variations du rayon ont été reliées aux propriétés du bois (micro-densitométrie) au niveau de leur déterminisme génétique. Ces deux caractères sont également liés au niveau phénotypique. En effet, des études ont mis en lien la mesure de micro-dendrométrie avec l'élaboration des propriétés [51, 67, 68]. Ce résultat nous semble majeur compte tenu de l'importance de la qualité du bois dans les programmes d'amélioration. L'analyse combinée de l'élaboration des propriétés du bois et de la réponse de la croissance à



l'environnement, pourrait permettre une meilleure compréhension des bases génétiques des propriétés du bois. Par ailleurs, une caractérisation des propriétés du bois a été engagée sur ce dispositif, dans le cadre d'un projet européen (Projet ERANET 34Joules). Cela va permettre de comparer de manière plus précise l'architecture génétique des propriétés du bois et celle de la croissance sur la base des cartes génétique haute résolution. Enfin, un projet a été soumis à l'ANR avec pour objectif d'identifier les gènes sous-jacents à la composition chimique du bois (comme le pourcentage de lignine) grâce au phénotypage et au génotypage de l'ensemble des individus présents sur le dispositif P10 ( $\approx 2500$ ). Les résultats de cette thèse sont de nature à amener des informations utiles à ce projet.

#### 7.2.4 Vers la sélection/prédiction génomique

Proposé en 2001 [69], le principe de la sélection génomique n'a été testé sur les arbres forestiers que récemment [70-72]. La réduction des coûts de génotypage liés aux développements des technologies haut débit [73], a permis l'accès à un nombre suffisant de marqueurs moléculaire pour couvrir le génome de manière efficace. Cette méthode a montré de bons résultats chez l'*Eucalyptus* [70] et *Pinus taeda* [71] avec des niveaux de précision supérieur à 0.8. Cependant, l'interaction génotype  $\times$  environnement reste un point clé dans la prédiction du phénotype grâce à l'information moléculaire. En effet, Resende *et al.* [72] ont montré sur *Pinus taeda* une diminution drastique de la précision ( $\approx 0.2$ ) des modèles de sélection génomique appliqués à différents environnements. Il serait donc intéressant de combiner ces approches de prédiction génomique avec des caractères de réponses à l'environnement comme abordé dans cette thèse. Dans ce cadre, le projet ANR WUETREE (démarré en 2014) vise à appliquer les modèles de sélection génomique sur des caractères comme le  $\delta^{13}\text{C}$  et à tester son utilisation dans le cadre du programme d'amélioration génétique de l'eucalyptus en République du Congo.





## Références

1. Thumma BR, Baltunis BS, Bell JC, Emebiri LC, Moran GF, Southerton SG: **Quantitative trait locus (QTL) analysis of growth and vegetative propagation traits in Eucalyptus nitens full-sib families.** *Tree Genetics & Genomes* 2010, **6**(6):877-889.
2. Thumma BR, Southerton SG, Bell JC, Owen JV, Henery ML, Moran GF: **Quantitative trait locus (QTL) analysis of wood quality traits in Eucalyptus nitens.** *Tree Genetics & Genomes* 2010, **6**(2):305-317.
3. Lerceteau E, Szmidt AE, Andersson B: **Detection of quantitative trait loci in Pinus sylvestris L. across years.** *Euphytica* 2001, **121**(2):117-122.
4. Emebiri LC, Devey ME, Matheson AC, Slee MU: **Age-related changes in the expression of QTLs for growth in radiata pine seedlings.** *TAG Theoretical and Applied Genetics* 1998, **97**(7):1053-1061.
5. Freeman JS, Whittock SP, Potts BM, Vaillancourt RE: **QTL influencing growth and wood properties in Eucalyptus globulus.** *Tree Genetics & Genomes* 2009, **5**(4):713-722.
6. Rae A, Pinel M, Bastien C, Sabatti M, Street N, Tucker J, Dixon C, Marron N, Dillen S, Taylor G: **QTL for yield in bioenergy Populus: identifying G×E interactions from growth at three contrasting sites.** *Tree Genetics & Genomes* 2008, **4**(1):97-112.
7. Rae AM, Street NR, Robinson KM, Harris N, Taylor G: **Five QTL hotspots for yield in short rotation coppice bioenergy poplar: The Poplar Biomass Loci.** *BMC Plant Biology* 2009, **9**.
8. Bundock PC, Potts BM, Vaillancourt RE: **Detection and stability of quantitative trait loci (QTL) in Eucalyptus globulus.** *Tree Genetics & Genomes* 2008, **4**(1):85-95.
9. Dillen SY, Storme V, Marron N, Bastien C, Neyrinck S, Steenackers M, Ceulemans R, Boerjan W: **Genomic regions involved in productivity of two interspecific poplar families in Europe. 1. Stem height, circumference and volume.** *Tree Genetics & Genomes* 2009, **5**(1):147-164.
10. Kullán AR, van Dyk M, Hefer C, Jones N, Kanzler A, Myburg A: **Genetic dissection of growth, wood basic density and gene expression in interspecific backcrosses of Eucalyptus grandis and E. urophylla.** *BMC Genetics* 2012, **13**(1):60.
11. Pelgas B, Bousquet J, Meirmans P, Ritland K, Isabel N: **QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments.** *BMC Genomics* 2011, **12**(1):145.
12. Freeman JS, Potts BM, Downes GM, Pilbeam D, Thavamanikumar S, Vaillancourt RE: **Stability of quantitative trait loci for growth and wood properties across multiple pedigrees and environments in Eucalyptus globulus.** *New Phytol* 2013, **198**(4):1121-1134.
13. Hudson CJ, Freeman JS, Kullán ARK, Petroli CD, Sansaloni CP, Kilian A, Detering F, Grattapaglia D, Potts BM, Myburg AA *et al*: **A reference linkage map for eucalyptus.** *BMC Genomics* 2012, **13**(1):240.
14. Petroli CD, Sansaloni CP, Carling J, Steane DA, Vaillancourt RE, Myburg AA, da Silva OB, Jr., Pappas GJ, Jr., Kilian A, Grattapaglia D: **Genomic Characterization of DArT Markers Based on High-Density Linkage Analysis and Physical Mapping to the Eucalyptus Genome.** *PLoS ONE* 2012, **7**(9):e44684.
15. Kullán ARK, van Dyk M, Jones N, Kanzler A, Bayley A, Myburg A: **High-density genetic linkage maps with over 2,400 sequence-anchored DArT markers for genetic dissection in an F2 pseudo-backcross of Eucalyptus grandis × E. urophylla.** *Tree Genetics & Genomes* 2012, **8**(1):163-175.
16. Ganai MW, Durstewitz G, Polley A, Bérard A, Buckler ES, Charcosset A, Clarke JD, Graner E-M, Hansen M, Joets J *et al*: **A Large Maize (Zea mays L.) SNP Genotyping Array: Development and Germplasm Genotyping, and Genetic Mapping to Compare with the B73 Reference Genome.** *PLoS ONE* 2011, **6**(12):e28334.
17. Felcher KJ, Coombs JJ, Massa AN, Hansey CN, Hamilton JP, Veilleux RE, Buell CR, Douches DS: **Integration of Two Diploid Potato Linkage Maps with the Potato Genome Sequence.** *PLoS ONE* 2012, **7**(4):e36347.



18. Wang Y, Sun S, Liu B, Wang H, Deng J, Liao Y, Wang Q, Cheng F, Wang X, Wu J: **A sequence-based genetic linkage map as a reference for Brassica rapa pseudochromosome assembly.** *BMC Genomics* 2011, **12**(1):239.
19. Ren Y, Zhao H, Kou Q, Jiang J, Guo S, Zhang H, Hou W, Zou X, Sun H, Gong G *et al*: **A High Resolution Genetic Map Anchoring Scaffolds of the Sequenced Watermelon Genome.** *PLoS ONE* 2012, **7**(1):e29453.
20. Grattapaglia D, Plomion C, Kirst M, Sederoff RR: **Genomics of growth traits in forest trees.** *Current opinion in plant biology* 2009, **12**(2):148-156.
21. Kole C: **Genome Mapping and Molecular Breeding in Plants: Forest trees**, vol. 7. Berlin: Springer; 2007.
22. Emebiri LC: **Detection and genetic mapping of quantitative trait loci influencing stem growth efficiency in radiata pine.** *Ph.D.*: Australian National University; 1997.
23. De Martonne E: **Une Nouvelle fonction climatologique. L'Indice d'aridité.** Paris: Impr. Gauthier-Villars; 1926.
24. Villar E: **Plasticité phénotypique et moléculaire de deux clones d'eucalyptus sous contrainte hydrique au champ.** *Ph.D.* Bordeaux: Université de Bordeaux I; 2011.
25. Cornelius J: **Heritabilities and additive genetic coefficients of variation in forest trees.** *Canadian Journal of Forest Research* 1994, **24**(2):372-379.
26. Bouvet JM, Vigneron P, Gouma R, Saya A: **Trends in variance and heritabilities with age for growth traits in Eucalyptus spacing experiments.** *Silvae genetica* 2003, **52**(3-4):121-133.
27. Rönnerberg-Wästljung AC, Glynn C, Weih M: **QTL analyses of drought tolerance and growth for a Salix dasyclados × Salix viminalis hybrid in contrasting water regimes.** *TAG Theoretical and Applied Genetics* 2005, **110**(3):537-549.
28. Marguerit E, Brendel O, Lebon E, Van Leeuwen C, Ollat N: **Rootstock control of scion transpiration and its acclimation to water deficit are controlled by different genes.** *New Phytol* 2012, **194**(2):416-429.
29. Schulze E-D, Turner NC, Nicolle D, Schumacher J: **Leaf and wood carbon isotope ratios, specific leaf areas and wood growth of Eucalyptus species across a rainfall gradient in Australia.** *Tree Physiology* 2006, **26**(4):479-492.
30. Zhang J, Fins L, Marshall JD: **Stable carbon isotope discrimination, photosynthetic gas exchange, and growth differences among western larch families.** *Tree Physiology* 1994, **14**(5):531-539.
31. Xu Z, Prasolova N, Lundkvist K, Beadle C, Leaman T: **Genetic variation in branchlet carbon and nitrogen isotope composition and nutrient concentration of 11-year-old hoop pine families in relation to tree growth in subtropical Australia.** *Forest Ecology and Management* 2003, **186**(1-3):359-371.
32. Monclus R, Dreyer E, Delmotte FM, Villar M, Delay D, Boudouresque E, Petit J-M, Marron N, Bréchet C, Brignolas F: **Productivity, leaf traits and carbon isotope discrimination in 29 Populus deltoides × P. nigra clones.** *New Phytol* 2005, **167**(1):53-62.
33. Monclus R, Dreyer E, Villar M, Delmotte FM, Delay D, Petit J-M, Barbaroux C, Le Thiec D, Bréchet C, Brignolas F: **Impact of drought on productivity and water use efficiency in 29 genotypes of Populus deltoides × Populus nigra.** *New Phytol* 2006, **169**(4):765-777.
34. Weih M, Ronnerberg-Wästljung AC, Glynn C: **Genetic basis of phenotypic correlations among growth traits in hybrid willow (Salix dasyclados × S-viminalis) grown under two water regimes.** *New Phytol* 2006, **170**(3):467-477.
35. Johnsen KH, Flanagan LB, Huber DA, Major JE: **Genetic variation in growth, carbon isotope discrimination, and foliar N concentration in Picea mariana: analyses from a half-diallel mating design using field-grown trees.** *Canadian Journal of Forest Research* 1999, **29**(11):1727-1735.
36. Monclus R, Leple J-C, Bastien C, Bert P-F, Villar M, Marron N, Brignolas F, Jorge V: **Integrating genome annotation and QTL position to identify candidate genes for productivity, architecture and water-use efficiency in Populus spp.** *BMC Plant Biology* 2012, **12**(1):173.



37. Nguyen-Queyrens A, Ferhi A, Loustau D, Guehl J-M: **Within-ring  $\delta^{13}\text{C}$  spatial variability and interannual variations in wood cellulose of two contrasting provenances of *Pinus pinaster*.** *Canadian Journal of Forest Research* 1998, **28**(5):766-773.
38. Brendel O, Pot D, Plomion C, Rozenberg P, Guehl JM: **Genetic parameters and QTL analysis of  $\delta^{13}\text{C}$  and ring width in maritime pine.** *Plant, Cell & Environment* 2002, **25**(8):945-953.
39. Prasolova NV, Xu Z, Farquhar GD, Saffigna PG, Dieters MJ: **Canopy carbon and oxygen isotope composition of 9-year-old hoop pine families in relation to seedling carbon isotope composition, growth, field growth performance, and canopy nitrogen concentration.** *Canadian Journal of Forest Research* 2001, **31**(4):673-681.
40. Jeon BW, Hwang J-U, Hwang Y, Song W-Y, Fu Y, Gu Y, Bao F, Cho D, Kwak JM, Yang Z *et al*: **The Arabidopsis Small G Protein ROP2 Is Activated by Light in Guard Cells and Inhibits Light-Induced Stomatal Opening.** *The Plant Cell Online* 2008, **20**(1):75-87.
41. Jung C, Seo JS, Han SW, Koo YJ, Kim CH, Song SI, Nahm BH, Choi YD, Cheong J-J: **Overexpression of AtMYB44 Enhances Stomatal Closure to Confer Abiotic Stress Tolerance in Transgenic Arabidopsis.** *Plant Physiol* 2008, **146**(2):623-635.
42. Kozlowski TT, Winget CH: **Diurnal and Seasonal Variation in Radii of Tree Stems.** *Ecology* 1964, **45**(1):149-155.
43. Herzog K, Häsler R, Thum R: **Diurnal changes in the radius of a subalpine Norway spruce stem: their relation to the sap flow and their use to estimate transpiration.** *Trees* 1995, **10**(2):94-101.
44. Offenthaler I, Hietz P, Richter H: **Wood diameter indicates diurnal and long-term patterns of xylem water potential in Norway spruce.** *Trees - Structure and Function* 2001, **15**(4):215-221.
45. Downes G, Beadle C, Worledge D: **Daily stem growth patterns in irrigated *Eucalyptus globulus* and *E. nitens* in relation to climate.** *Trees* 1999, **14**(2):102-111.
46. Deslauriers A, Morin H, Urbinati C, Carrer M: **Daily weather response of balsam fir (*Abies balsamea* (L.) Mill.) stem radius increment from dendrometer analysis in the boreal forests of Québec (Canada).** *Trees* 2003, **17**(6):477-484.
47. Bouriaud O, Leban J-M, Bert D, Deleuze C: **Intra-annual variations in climate influence growth and wood density of Norway spruce.** *Tree Physiology* 2005, **25**(6):651-660.
48. Zweifel R, Item H, Häsler R: **Stem radius changes and their relation to stored water in stems of young Norway spruce trees.** *Trees* 2000, **15**(1):50-57.
49. Zweifel R, Item H, Häsler R: **Link between diurnal stem radius changes and tree water relations.** *Tree Physiology* 2001, **21**(12-13):869-877.
50. Cochard H, Forestier S, Améglio T: **A new validation of the Scholander pressure chamber technique based on stem diameter variations.** *Journal of Experimental Botany* 2001, **52**(359):1361-1365.
51. Drew DM, Downes GM: **The use of precision dendrometers in research on daily stem size and wood property variation: A review.** *Dendrochronologia* 2009, **27**(2):159-172.
52. Deslauriers A, Anfodillo T, Rossi S, Carraro V: **Using simple causal modeling to understand how water and temperature affect daily stem radial variation in trees.** *Tree Physiology* 2007, **27**(8):1125-1136.
53. Zweifel R, Zimmermann L, Zeugin F, Newbery DM: **Intra-annual radial growth and water relations of trees: implications towards a growth mechanism.** *Journal of Experimental Botany* 2006, **57**(6):1445-1459.
54. Drew D, Downes G, Grzeskowiak V, Naidoo T: **Differences in daily stem size variation and growth in two hybrid eucalypt clones.** *Trees - Structure and Function* 2009, **23**(3):585-595.
55. Deslauriers A, Rossi S, Anfodillo T: **Dendrometer and intra-annual tree growth: What kind of information can be inferred?** *Dendrochronologia* 2007, **25**(2):113-124.
56. Zweifel R, Zimmermann L, Newbery DM: **Modeling tree water deficit from microclimate: an approach to quantifying drought stress.** *Tree Physiology* 2005, **25**(2):147-156.



57. Wu R, Lin M: **Functional mapping : how to map and study the genetic architecture of dynamic complex traits.** *Nat Rev Genet* 2006, **7**(3):229-237.
58. Groenen MAM, Wahlberg P, Foglio M, Cheng HH, Megens H-J, Crooijmans RPMA, Besnier F, Lathrop M, Muir WM, Wong GK-S *et al*: **A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate.** *Genome Res* 2009, **19**(3):510-519.
59. Giraut L, Falque M, Drouaud J, Pereira L, Martin OC, Mézard C: **Genome-Wide Crossover Distribution in *Arabidopsis thaliana* Meiosis Reveals Sex-Specific Patterns along Chromosomes.** *PLoS Genet* 2011, **7**(11):e1002354.
60. Henderson IR: **Control of meiotic recombination frequency in plant genomes.** *Current opinion in plant biology* 2012, **15**(5):556-561.
61. McVean GAT, Myers SR, Hunt S, Deloukas P, Bentley DR, Donnelly P: **The Fine-Scale Structure of Recombination Rate Variation in the Human Genome.** *Science* 2004, **304**(5670):581-584.
62. Nachman MW: **Variation in recombination rate across the genome: evidence and implications.** *Current Opinion in Genetics & Development* 2002, **12**(6):657-663.
63. McClosky B, Tanksley SD: **The impact of recombination on short-term selection gain in plant breeding experiments.** *Theor Appl Genet* 2013, **126**(9):2299-2312.
64. Ma CX, Wu R, Casella G: **FunMap: functional mapping of complex traits.** *Bioinformatics* 2004, **20**(11):1808-1811.
65. Li Y, Wu R: **Functional mapping of growth and development.** *Biological Reviews* 2010, **85**(2):207-216.
66. Xiong H, Goulding EH, Carlson EJ, Tecott LH, McCulloch CE, Sen S: **A Flexible Estimating Equations Approach for Mapping Function-Valued Traits.** *Genetics* 2011, **189**(1):305-316.
67. Wimmer R, Downes GM, Evans R: **Temporal variation of microfibril angle in *Eucalyptus nitens* grown in different irrigation regimes.** *Tree Physiology* 2002, **22**(7):449-457.
68. Wimmer R, Downes GM, Evans R: **High-resolution analysis of radial growth and wood density in *Eucalyptus nitens*, grown under different irrigation regimes.** *Ann For Sci* 2002, **59**(5-6):519-524.
69. Meuwissen THE, Hayes BJ, Goddard ME: **Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps.** *Genetics* 2001, **157**(4):1819-1829.
70. Resende MDV, Resende MFR, Sansaloni CP, Petroli CD, Missiaggia AA, Aguiar AM, Abad JM, Takahashi EK, Rosado AM, Faria DA *et al*: **Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits in forest trees.** *New Phytol* 2012, **194**(1):116-128.
71. Resende MFR, Muñoz P, Resende MDV, Garrick DJ, Fernando RL, Davis JM, Jokela EJ, Martin TA, Peter GF, Kirst M: **Accuracy of Genomic Selection Methods in a Standard Data Set of Loblolly Pine (*Pinus taeda* L.).** *Genetics* 2012, **190**(4):1503-1510.
72. Resende MFR, Muñoz P, Acosta JJ, Peter GF, Davis JM, Grattapaglia D, Resende MDV, Kirst M: **Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments.** *New Phytol* 2012, **193**(3):617-624.
73. Appleby N, Edwards D, Batley J: **New Technologies for Ultra-High Throughput Genotyping in Plants.** In: *Plant Genomics*. Edited by Gustafson JP, Langridge P, Somers DJ, vol. 513: Humana Press; 2009: 19-39.





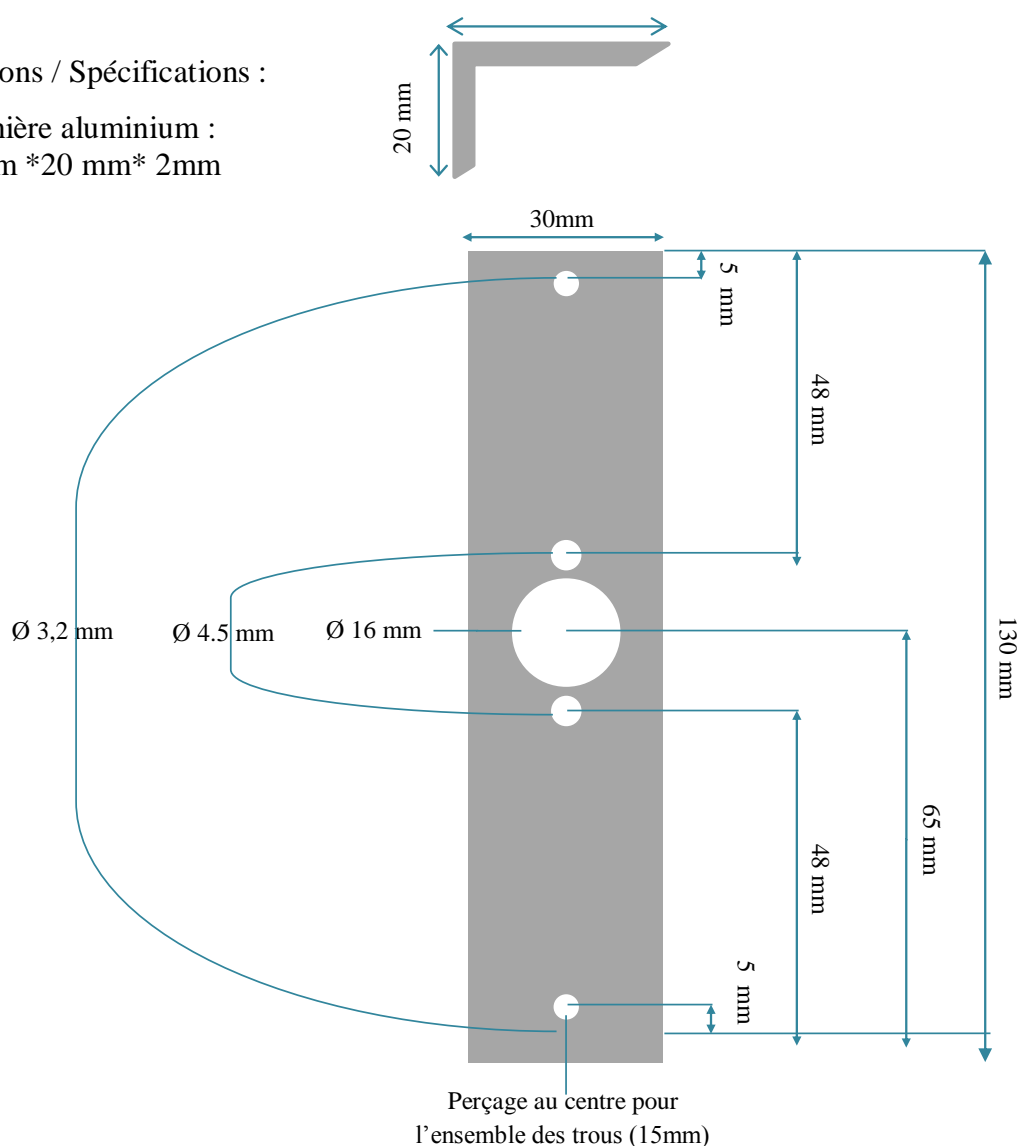
# Annexes

Annexe 1 : Conception du support de capteur pour le dispositif P10.....	279
Annexe 2 : Guide des préparatifs effectués au CRDPI (pointe-noire) en amont de l'installation du dispositif P10.....	280
Annexe 3 : Programme de mesure pour la centrale du plateau et pour la centrale météo. ....	287
Annexe 4 : Protocole d'extraction d'adn optimisé en plaque 96 avec le kit ademtech. ....	296

## Annexe 1 : Conception du support de capteur pour le dispositif P10.

Dimensions / Spécifications :

Cornière aluminium :  
30mm \*20 mm\* 2mm



Le côté de la cornière mesurant 20 mm ne fait l'objet d'aucune intervention.



Annexe 2 : Guide des préparatifs effectués au CRDPI (Pointe-Noire) en amont de l'installation du dispositif P10



# Planification des activités ABIOTEN au CRDPI

Avril - Mai 2011

Jérôme Bartholomé





## I. Protocole pour la Préparation des câbles

Il est prévu pour André Mabiala deux semaines pour la préparation des câbles avant l'arrivée de J. Bartholomé le 29 avril 2011. Le matériel de découpe est disponible au Congo, Les câbles ont été envoyés par bateau pour le 2 avril (réceptionné à Pointe Noire par SDV).

### A. Différentes étapes de préparation des câbles

Remarque : Deux types de câbles ont été envoyés : **câbles noirs** et **câbles oranges**. Les câbles qui doivent être préparés sont uniquement les **câbles noirs**.

Faire les étapes suivantes par série de câbles : voir tableau ci-dessous

- Découper tous les câbles de même longueur
- Dénuder les extrémités (explications ci-dessous, chapitre 3)
- Rassembler les câbles préparés et indiquer leur longueur sur des étiquettes

Ex : Couper tous les câbles de 9 mètres d'une série puis dénuder les extrémités et rassembler les câbles par paquet en indiquant leur longueur sur une étiquette. Ensuite faire pareil avec une autre longueur de câble de la même série. Quand toutes les longueurs de câble de la première série sont faites, passer à la deuxième série.

**Ne pas mélanger les différentes longueurs de câble.**

### B. Découpe des câbles

#### Série 1

Longueur de câble (m)	9	11	14	16	19	21	24
nombre de câbles	11	12	22	11	20	11	10

#### Série 2

Longueur de câble (m)	9	11	14	16	19	21	24
nombre de câbles	11	11	23	12	20	11	9

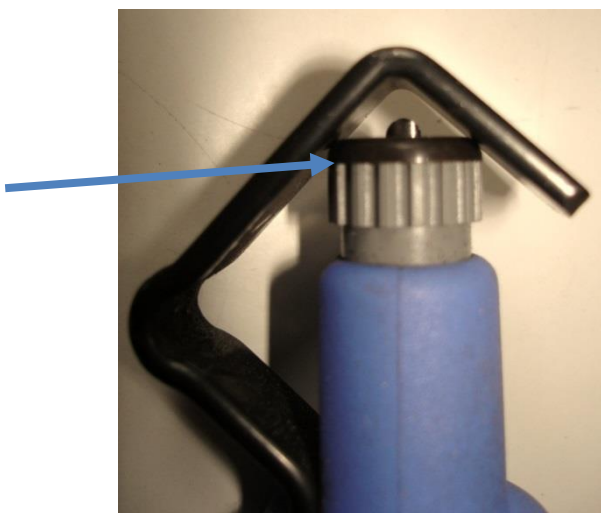
La découpe des câbles s'effectuera avec la pince coupante aux poignées rouges (disponible dans la malle ABIOPEN). Les câbles sont enroulés et déposés dans un carton avec la longueur indiquée sur le carton.



### C. Dénuder la gaine plastique des extrémités des câbles

Régler la profondeur de découpe avec la molette afin de ne pas couper la gaine métallique. Faire des essais sur un bout de câble avant d'appliquer en série.

Molette de réglage



Dénuder la gaine plastique des deux extrémités de chaque câble. Il y a une petite extrémité (dénudée sur **8 cm** de long) et une grande extrémité (dénudée sur **80 cm** de long), pour chaque câble. La gaine métallique ne doit pas être abîmée et/ou enlevée à ce stade.

Grande extrémité

Petite extrémité



## D. Dénuder la gaine métallique des extrémités des câbles

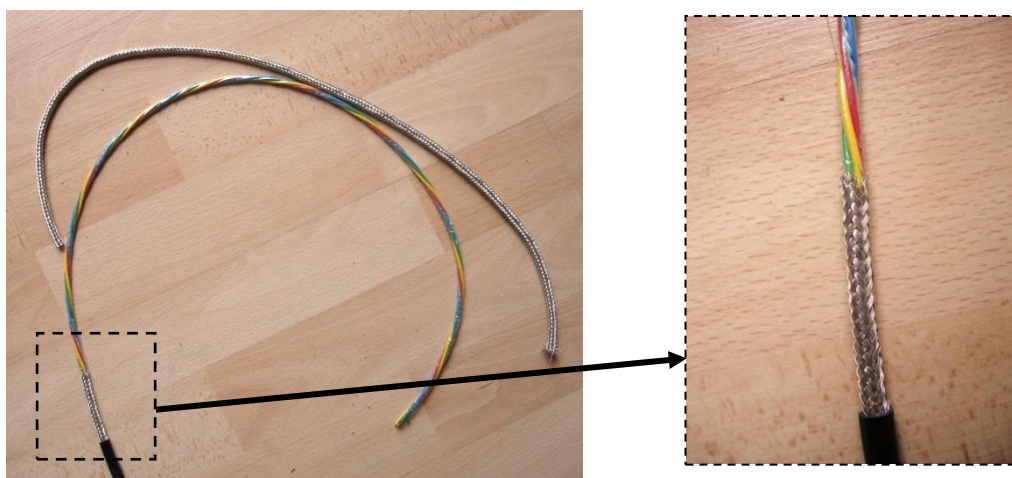
Ce travail de dénudage doit être effectué avec la pince coupante aux poignées bleues.

- **La petite extrémité** doit être dénudée **sur 7-8 cm** (environ 4 doigts). Il faut enlever la **totalité** de la gaine métallique ainsi que le fil bleu (voir photo ci-dessous).



- **La grande extrémité** doit être dégainée sur **80 cm** (environ la longueur d'un bras, de la main jusqu'à l'épaule)

Enlever la gaine métallique **en laissant 5 cm (3 doigts) à la base** (voir photo ci-dessous)



Ensuite enlever le plastique ainsi que le fil bleu, toujours avec la pince aux poignées bleues.

## E. Classement des câbles dénudés

Chaque classe de longueur de câble doit être enroulée et liée avec un serre-fil (prendre les serre-fils blancs). Une étiquette sur le serre-fil indiquera la longueur des câbles attachés. Chaque classe de longueur sera entreposée dans le carton utilisé précédemment.



## II. Préparation du terrain

Démarrage des travaux par la cabane du gardien et le grillage la troisième semaine d'avril.  
Le 18 avril 2011. L. Batekidila et P. Perrin sont en charge de ces travaux.

Il y a un délai de 15 jours pour la livraison de poteaux pour le grillage (voir ci-dessous)

**SNBI S.A**

SOCIETE NOUVELLE DES BOIS IMPREGNES

Pointe-Noire, le 23 Mars 2011

**FACTURE PROFORMA N° 005/SNBI/11**

Client : CRDPI  
Tél : 06 970 87 89  
Pointe-Noire

Marchandise : Poteaux en bois  
d'eucalyptus traités au CCA pour  
supports de lignes électriques types SNE

**Spécifications Technique****Poteaux de 3 mètres**

Diamètre au sommet	:	12 cm
Diamètre à la base	:	14 cm
Volume	:	0,039 m3

**Quantités & Coûts usine HT**

140 poteaux de 3 m x 8 276 F.cfa = 1 158 667 F.cfa

**Montant total HT** **1 158 667 F.cfa**

TVA 18% 208 560 F.cfa  
CA 5% 10 428 F.cfa

**Montant total HT** **1 377 655 F.cfa**

**Délai de livraison : Deux Semaines.**

Arrête la présente facture pro forma à la somme de : **UN MILLION TROIS CENT SOIXANTE DIX SEPT MILLE SIX CENT SOIXANTE CINQ FRANCS CFA.**

Le Conseiller Technique de  
l'Administrateur Général

Désiré TCHITEMBO



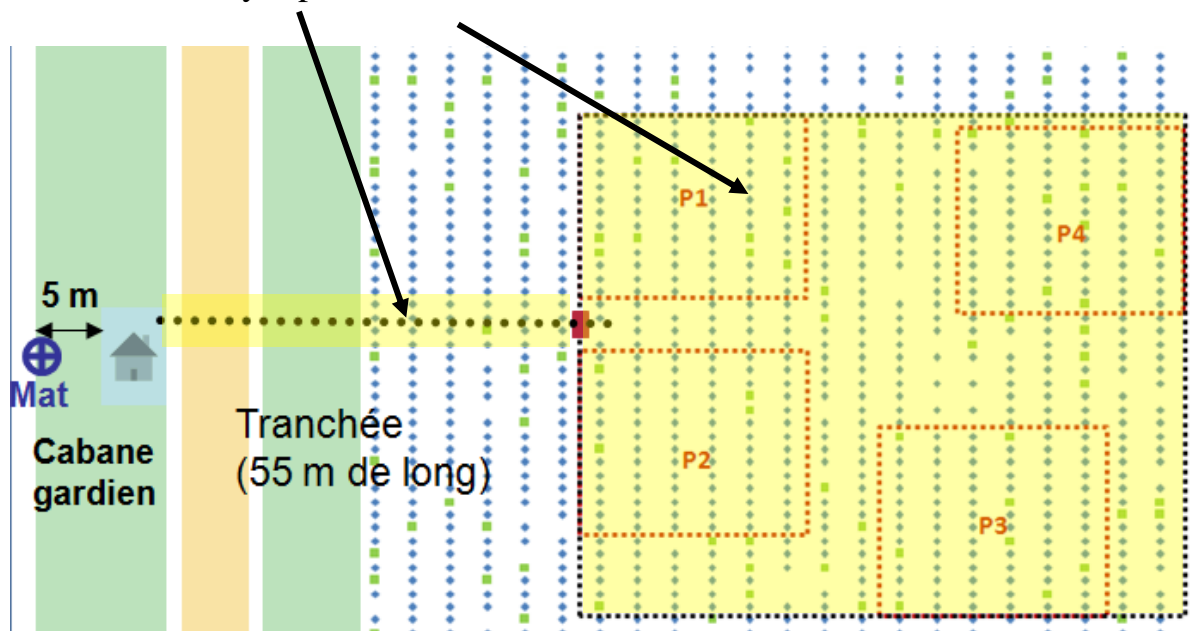
Les marchandises vendues ne sont ni reprises, ni échangées, ni remboursées ; elles restent la propriété du vendeur jusqu'à paiement intégral du prix. Les risques afférents aux dites marchandises sont transférés à l'acheteur dès livraison

SOCIETE NOUVELLE DES BOIS IMPREGNES S.A (SNBI)  
AV. CHARLES DE GAULLE TOUR NAVOIRE ETAGE BP 854 POINTE-NOIRE  
E-mail : [snbi\\_cg@yahoo.com](mailto:snbi_cg@yahoo.com) / BUI : 800011000020153  
TEL : (242) 2 94 65 04 FAX : (242) 2 94 53 04



Nettoyer le terrain à l'intérieur du grillage ainsi que le passage qui relie l'entrée du grillage jusqu'à la cabane du gardien (pour le passage des câbles d'alimentation) **avant la fin avril**.  
**Lors du nettoyage faire attention aux arbres.**

Zone à nettoyer pour fin avril



### III. Gardiennage

Prévoir le début du gardiennage lors de la pose du grillage.

### IV. Inventaire du matériel présent au Congo

Faire un inventaire du matériel présent dans la malle ABIOPEN pour **le 15 avril**.

Matériel	Quantité/ état	Matériel	Quantité/ état
Pince coupe câble 162 mm		Collier desserrable	
Outil à dégainer		Collier standard	
Pince coupante (bleu)		Marqueur permanent	
Pince à dénuder 15-1mm (verte)		Etiquette repérage pour câble	
Jeu de 10 tournevis		Stylo feutre	
3 limes 15 cm		Scotch	
Jeu de 9 clés		Perceuse	
Clef à molette 12 pouces		Boitier + Centrale	
Serre fils non desserrable		Sonde TDR	



## V. Calendrier prévisionnel : Avril Mai 2011

	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche
Avril	18 Découpe et préparation des câbles	19 Découpe et préparation des câbles	20 Découpe et préparation des câbles	21 Découpe et préparation des câbles	22 Découpe et préparation des câbles	23	24
	25 Découpe et préparation des câbles	26 Découpe et préparation des câbles	27 Découpe et préparation des câbles	28 Découpe et préparation des câbles	29 Arrivée Jérôme	30 Câblage des centrales	1 Câblage des centrales
					Câblage des centrales		
MAI	2 Câblage	3 Câblage	4 Câblage	5 Câblage	6 Câblage	7 Câblage	8 Câblage
				Installation terrain	Installation terrain	Installation terrain	Installation terrain
				Arrivée Jean-Marc et Didier			
	9 Installation terrain	10 Installation terrain	11 Installation terrain	12 Installation terrain	13 Installation terrain	14 Installation terrain	15 Installation terrain
		Arrivée Régis			Départ Jean Marc et Didier		
	16 Installation terrain	17 Installation terrain	18 Installation terrain	19 Installation terrain	20 Installation terrain	21 Installation terrain	22 Installation terrain
					Départ Régis		
	23 Installation terrain	24 Vérification	26 Vérification	26 Vérification	27 Départ Jérôme	28	29





## Annexe 3 : Programme de mesure pour la centrale du plateau et pour la centrale météo.

### Programme de la centrale du plateau 1 (P1)

```
.....
' CR1000 _ packbus 2
' PROGRAMME manip ABIOTEN CONGO
' PARCELLE 1
'v2
' Regis.Burlett@u-bordeaux1.fr
' jerome.bartholome@pierroton.inra.fr
.....

' New in version 1.1
'      modification d'un nom de capteur ! L'arbre 12_29 est mort, et il a ete remplace DES LE DEBUT par
le 12_30, mais pas modifie dans la v 1.0

' New in version 1.3
'      * Passage une mesure toutes les 30 sec et moyennes a 5 min

'New in version 2
'      Elimination de l'envoi du fichier

'Declare Variables and Units

Dim LCount_17
Dim LCount_19
Public BattV_CP1
Public TRHData(2)
Public T_SOIL_CP1_15cm
Public T_SOIL_CP1_100cm
Public PA_uS_1_CP1
Public SWC_15cm_1_CP1
Public PA_uS_2_CP1
Public SWC_50cm_2_CP1
Public PA_uS_3_CP1
Public SWC_100cm_CP1
Public PA_uS_4_CP1
Public SWC_200cm_CP1

Public timestamp_senddata As String *20
Public LastFileName_CP1 As String *40
Public YYYY_doy_senddata As String * 20
Public Filename_day_CP1_tosend As String * 35
'Public Filename_day_CP1_rename As String * 35

Dim time_array (9)

Public variable
Public Result_sendCP1

Public outstat_CP1 As Long 'variable that hold the status of the daily file construction

Public year_yesterday
Public doy_yesterday
```





Public Vref\_MUX\_CA1  
Public Vref\_MUX\_CB1

Public SEVolt\_MUXA(33)  
Public SEVolt\_MUXB(33)

Alias SEVolt\_MUXA(1)= C\_X8\_Y23  
Alias SEVolt\_MUXA(2)= C\_X8\_Y25  
Alias SEVolt\_MUXA(3)= C\_X8\_Y27  
Alias SEVolt\_MUXA(4)= C\_X8\_Y28  
Alias SEVolt\_MUXA(5)= C\_X8\_Y29  
Alias SEVolt\_MUXA(6)= C\_X9\_Y25  
Alias SEVolt\_MUXA(7)= C\_X9\_Y26  
Alias SEVolt\_MUXA(8)= C\_X9\_Y27  
Alias SEVolt\_MUXA(9)= C\_X9\_Y29  
Alias SEVolt\_MUXA(10)= C\_X10\_Y24  
Alias SEVolt\_MUXA(11)= C\_X10\_Y25  
Alias SEVolt\_MUXA(12)= C\_X10\_Y27  
Alias SEVolt\_MUXA(13)= C\_X10\_Y28  
Alias SEVolt\_MUXA(14)= C\_X10\_Y29  
Alias SEVolt\_MUXA(15)= C\_X11\_Y26  
Alias SEVolt\_MUXA(16)= C\_X11\_Y27  
Alias SEVolt\_MUXA(17)= C\_X11\_Y28  
Alias SEVolt\_MUXA(18)= C\_X11\_Y29  
Alias SEVolt\_MUXA(19)= C\_X12\_Y24  
Alias SEVolt\_MUXA(20)= C\_X12\_Y25  
Alias SEVolt\_MUXA(21)= C\_X12\_Y26  
Alias SEVolt\_MUXA(22)= C\_X12\_Y27  
Alias SEVolt\_MUXA(23)= C\_X12\_Y28  
Alias SEVolt\_MUXA(24)= C\_X12\_Y30  
Alias SEVolt\_MUXA(25)= C\_X13\_Y24  
Alias SEVolt\_MUXA(26)= C\_X13\_Y26  
Alias SEVolt\_MUXA(27)= C\_X13\_Y27  
Alias SEVolt\_MUXA(28)= C\_X13\_Y28  
Alias SEVolt\_MUXA(29)= C\_X8\_Y26

Alias SEVolt\_MUXB(1)= C\_X8\_Y20  
Alias SEVolt\_MUXB(2)= C\_X8\_Y21  
Alias SEVolt\_MUXB(3)= C\_X8\_Y22  
Alias SEVolt\_MUXB(4)= C\_X9\_Y18  
Alias SEVolt\_MUXB(5)= C\_X9\_Y21  
Alias SEVolt\_MUXB(6)= C\_X9\_Y22  
Alias SEVolt\_MUXB(7)= C\_X10\_Y16  
Alias SEVolt\_MUXB(8)= C\_X9\_Y19  
Alias SEVolt\_MUXB(9)= C\_X10\_Y18  
Alias SEVolt\_MUXB(10)= C\_X10\_Y19  
Alias SEVolt\_MUXB(11)= C\_X10\_Y20  
Alias SEVolt\_MUXB(12)= C\_X10\_Y21  
Alias SEVolt\_MUXB(13)= C\_X10\_Y22  
Alias SEVolt\_MUXB(14)= C\_X10\_Y23  
Alias SEVolt\_MUXB(15)= C\_X11\_Y17  
Alias SEVolt\_MUXB(16)= C\_X11\_Y18  
Alias SEVolt\_MUXB(17)= C\_X11\_Y19  
Alias SEVolt\_MUXB(18)= C\_X11\_Y22  
Alias SEVolt\_MUXB(19)= C\_X12\_Y17  
Alias SEVolt\_MUXB(20)= C\_X12\_Y19  
Alias SEVolt\_MUXB(21)= C\_X12\_Y20  
Alias SEVolt\_MUXB(22)= C\_X12\_Y22  
Alias SEVolt\_MUXB(23)= C\_X12\_Y23



Alias SEVolt\_MUXB(24)= C\_X13\_Y18  
Alias SEVolt\_MUXB(25)= C\_X13\_Y19  
Alias SEVolt\_MUXB(26)= C\_X13\_Y20  
Alias SEVolt\_MUXB(27)= C\_X13\_Y22  
Alias SEVolt\_MUXB(28)= C\_X13\_Y23  
Alias SEVolt\_MUXB(29)= C\_TP1\_1  
Alias SEVolt\_MUXB(30)= C\_TP1\_2  
Alias SEVolt\_MUXB(31)= C\_X9\_Y23

Alias TRHData(1)=AirTC\_CP1  
Alias TRHData(2)=RH\_CP1

Units BattV\_CP1=Volts  
Units T\_SOIL\_CP1\_15cm=Deg C  
Units T\_SOIL\_CP1\_100cm=Deg C  
Units PA\_uS\_1\_CP1=uSec  
Units PA\_uS\_2\_CP1=uSec  
Units PA\_uS\_3\_CP1=uSec  
Units PA\_uS\_4\_CP1=uSec  
Units SEVolt\_MUXA=mV  
Units SEVolt\_MUXB=mV  
Units AirTC\_CP1=Deg C  
Units RH\_CP1=%

'Define Data Tables

DataTable(DataTable\_CP1,True,-1)  
  DataInterval(0,5,Min,10)  
' CardOut(0,-1)  
TableFile ("CRD:DataTable\_CP1\_",9,-1,0,1440,Min,outstat\_CP1,LastFileName\_CP1)  
Average (1,BattV\_CP1,FP2,False)  
Average(1,AirTC\_CP1,FP2,False)  
Sample(1,RH\_CP1,FP2)  
Average(1,T\_SOIL\_CP1\_15cm,FP2,False)  
Average(1,T\_SOIL\_CP1\_100cm,FP2,False)  
Average(1,SWC\_15cm\_1\_CP1,FP2,False)  
Average(1,SWC\_50cm\_2\_CP1,FP2,False)  
Average(1,SWC\_100cm\_CP1,FP2,False)  
Average(1,SWC\_200cm\_CP1,FP2,False)  
Average(1,Vref\_MUX\_CA1,FP2,False)  
Average(1,Vref\_MUX\_CB1,FP2,False)  
Average(1,SEVolt\_MUXA(1),UINT2,False)  
Average(1,SEVolt\_MUXA(2),UINT2,False)  
Average(1,SEVolt\_MUXA(3),UINT2,False)  
Average(1,SEVolt\_MUXA(4),UINT2,False)  
Average(1,SEVolt\_MUXA(5),UINT2,False)  
Average(1,SEVolt\_MUXA(6),UINT2,False)  
Average(1,SEVolt\_MUXA(7),UINT2,False)  
Average(1,SEVolt\_MUXA(8),UINT2,False)  
Average(1,SEVolt\_MUXA(9),UINT2,False)  
Average(1,SEVolt\_MUXA(10),UINT2,False)  
Average(1,SEVolt\_MUXA(11),UINT2,False)  
Average(1,SEVolt\_MUXA(12),UINT2,False)  
Average(1,SEVolt\_MUXA(13),UINT2,False)  
Average(1,SEVolt\_MUXA(14),UINT2,False)  
Average(1,SEVolt\_MUXA(15),UINT2,False)  
Average(1,SEVolt\_MUXA(16),UINT2,False)  
Average(1,SEVolt\_MUXA(17),UINT2,False)  
Average(1,SEVolt\_MUXA(18),UINT2,False)  
Average(1,SEVolt\_MUXA(19),UINT2,False)  
Average(1,SEVolt\_MUXA(20),UINT2,False)



```
Average(1,SEVlt_MUXA(21),UINT2,False)
Average(1,SEVlt_MUXA(22),UINT2,False)
Average(1,SEVlt_MUXA(23),UINT2,False)
Average(1,SEVlt_MUXA(24),UINT2,False)
Average(1,SEVlt_MUXA(25),UINT2,False)
Average(1,SEVlt_MUXA(26),UINT2,False)
Average(1,SEVlt_MUXA(27),UINT2,False)
Average(1,SEVlt_MUXA(28),UINT2,False)
Average(1,SEVlt_MUXA(29),UINT2,False)
' Average(1,SEVlt_MUXA(30),UINT2,False)
' Average(1,SEVlt_MUXA(31),UINT2,False)
' Average(1,SEVlt(32),UINT2,False)
```

```
Average(1,SEVlt_MUXB(1),UINT2,False)
Average(1,SEVlt_MUXB(2),UINT2,False)
Average(1,SEVlt_MUXB(3),UINT2,False)
Average(1,SEVlt_MUXB(4),UINT2,False)
Average(1,SEVlt_MUXB(5),UINT2,False)
Average(1,SEVlt_MUXB(6),UINT2,False)
Average(1,SEVlt_MUXB(7),UINT2,False)
Average(1,SEVlt_MUXB(8),UINT2,False)
Average(1,SEVlt_MUXB(9),UINT2,False)
Average(1,SEVlt_MUXB(10),UINT2,False)
Average(1,SEVlt_MUXB(11),UINT2,False)
Average(1,SEVlt_MUXB(12),UINT2,False)
Average(1,SEVlt_MUXB(13),UINT2,False)
Average(1,SEVlt_MUXB(14),UINT2,False)
Average(1,SEVlt_MUXB(15),UINT2,False)
Average(1,SEVlt_MUXB(16),UINT2,False)
Average(1,SEVlt_MUXB(17),UINT2,False)
Average(1,SEVlt_MUXB(18),UINT2,False)
Average(1,SEVlt_MUXB(19),UINT2,False)
Average(1,SEVlt_MUXB(20),UINT2,False)
Average(1,SEVlt_MUXB(21),UINT2,False)
Average(1,SEVlt_MUXB(22),UINT2,False)
Average(1,SEVlt_MUXB(23),UINT2,False)
Average(1,SEVlt_MUXB(24),UINT2,False)
Average(1,SEVlt_MUXB(25),UINT2,False)
Average(1,SEVlt_MUXB(26),UINT2,False)
Average(1,SEVlt_MUXB(27),UINT2,False)
Average(1,SEVlt_MUXB(28),UINT2,False)
Average(1,SEVlt_MUXB(29),UINT2,False)
Average(1,SEVlt_MUXB(30),UINT2,False)
Average(1,SEVlt_MUXB(31),UINT2,False)
' Average(1,SEVlt_MUXB(32),UINT2,False)
EndTable
```

'Main Program

BeginProg

' SetStatus ("USRdrivesize", "2350000")

Scan(30,Sec,1,0)

RealTime(time\_array)

'Default Datalogger Battery Voltage measurement BattV

Battery(BattV\_CP1)

'CS215 Temperature & Relative Humidity Sensor measurements AirTC and RH

SDI12Recorder(TRHDData(),5,"0","M!",1,0)

'107 Temperature Probe (CSL) measurement T\_SOIL\_15cm

Therm107(T\_SOIL\_CP1\_15cm,1,9,Vx3,0,\_50Hz,1,0)



'107 Temperature Probe (CSL) measurement T\_SOIL45cm  
Therm107(T\_SOIL\_CP1\_100cm,1,10,Vx2,0,\_50Hz,1,0)

\*\*\*\*\*  
'CS616 Water Content Reflectometer measurements SWC\_1 and PA\_uS\_1\_CP1  
If IfTime(0,1,Hr) Then  
  CS616(PA\_uS\_1\_CP1,1,13,6,1,1,0)  
  SWC\_15cm\_1\_CP1=-0.0663+(-0.0063\*PA\_uS\_1\_CP1)+(0.0007\*PA\_uS\_1\_CP1^2)  
'EndIf

'CS616 Water Content Reflectometer measurements SWC\_2 and PA\_uS\_2\_CP1  
If IfTime(0,1,Hr) Then  
  CS616(PA\_uS\_2\_CP1,1,14,6,1,1,0)  
  SWC\_50cm\_2\_CP1=-0.0663+(-0.0063\*PA\_uS\_2\_CP1)+(0.0007\*PA\_uS\_2\_CP1^2)  
'EndIf

'CS616 Water Content Reflectometer measurements SWC\_3 and PA\_uS\_3\_CP1  
If IfTime(0,1,Hr) Then  
  CS616(PA\_uS\_3\_CP1,1,15,7,1,1,0)  
  SWC\_100cm\_CP1=-0.0663+(-0.0063\*PA\_uS\_3\_CP1)+(0.0007\*PA\_uS\_3\_CP1^2)  
'EndIf

'CS616 Water Content Reflectometer measurements SWC\_4 and PA\_uS\_4\_CP1  
If IfTime(0,1,Hr) Then  
  CS616(PA\_uS\_4\_CP1,1,16,7,1,1,0)  
  SWC\_200cm\_CP1=-0.0663+(-0.0063\*PA\_uS\_4\_CP1)+(0.0007\*PA\_uS\_4\_CP1^2)  
'EndIf

Delay(0,150,mSec)  
LCount\_17=1  
SubScan(0,mSec,11)

\*\*\*\*\*  
'ENABLE MUX  
PortSet (9 ,1 )

'Turn AM16/32 Multiplexer CA1 On  
PortSet(1,1)  
Delay(0,150,mSec)  
'Switch to next AM16/32 Multiplexer channel  
PulsePort(2,10000)

'Generic Single Ended Voltage measurements SEVolt() on the AM16/32 Multiplexer:  
VoltSe(SEVolt\_MUXA(LCount\_17),1,mV2500,1,False,0,\_50Hz,1,0)  
'  Delay(0,350,mSec)  
LCount\_17=LCount\_17+1  
VoltSe(SEVolt\_MUXA(LCount\_17),1,mV2500,2,False,0,\_50Hz,1,0)  
'  Delay(0,350,mSec)  
LCount\_17=LCount\_17+1  
VoltSe(SEVolt\_MUXA(LCount\_17),1,mV2500,3,True,0,\_50Hz,1,0)  
LCount\_17=LCount\_17+1  
NextSubScan

PortSet(1,0)  
Delay(0,150,mSec)  
' MESURE DU VOLTAGE 0 LA BORNE DES CAPTEURS DU MUX CA1  
  
VoltSe (Vref\_MUX\_CA1,1,mV2500,7,False,0,\_50Hz,1,0,0)



```
'calcul du ratio , et multiplication par 10000 pour eviter de stocker le 0 avant la virgule
For LCount_17 = 1 To 33
  SEVolt_MUXA(LCount_17)= 10000* SEVolt_MUXA(LCount_17) / Vref_MUX_CA1
Next

.....

'Turn AM16/32 Multiplexer CB1 On
PortSet(3,1)

Delay(0,150,mSec)

LCount_19=1
SubScan(0,uSec,11)

'Switch to next AM16/32 Multiplexer channel
PulsePort(4,10000)
'Generic Single Ended Voltage measurements SEVolt_MUXB() on the AM16/32 Multiplexer:

VoltSe(SEVolt_MUXB(LCount_19),1,mV2500,4,False,0,_50Hz,1,0)
LCount_19=LCount_19+1
'Generic Single Ended Voltage measurements SEVolt_MUXB() on the AM16/32 Multiplexer:

VoltSe(SEVolt_MUXB(LCount_19),1,mV2500,5,False,0,_50Hz,1,0)
LCount_19=LCount_19+1

"Generic Single Ended Voltage measurements SEVolt_MUXB() on the AM16/32 Multiplexer:
VoltSe(SEVolt_MUXB(LCount_19),1,mV2500,6,True,0,_50Hz,1,0)
LCount_19=LCount_19+1

NextSubScan

'Turn AM16/32 Multiplexer Off
PortSet(3,0)

' MESURE DU VOLTAGE 0 LA BORNE DES CAPTEURS DU MUX A6

VoltSe (Vref_MUX_CB1,1,mV2500,8,False,0,_50Hz,1,0,0)

'calcul du ratio , et multiplication par 10000 pour eviter de stocker le 0 avant la virgule
For LCount_17 = 1 To 33
  SEVolt_MUXB(LCount_17)= 10000* SEVolt_MUXB(LCount_17) / Vref_MUX_CB1

Next
.....
'DISABLE MUX
PortSet(9,0)

CallTable(DataTable_CPI)

'-----
' THIS SET OF INSTRUCTION DEFINE THE FILENAME TO SEND

If TimeIntoInterval (10,1440,Min) Then
  'Isolate only the date froml the timestamp in YYYY-MM-DD
  If time_array > 1 Then
    doy_yesterday = time_array(9) - 1
```



```
year_yesterday = time_array(1)
Else
  doy_yesterday = 365
  year_yesterday = time_array(1) - 1
EndIf

YYYY_doy_senddata = year_yesterday + "_" + doy_yesterday

Filename_day_CP1_tosend = "CRD:ABIOGEN_CON_P1_" & YYYY_doy_senddata + ".dat"

FileRename (LastFileName_CP1,Filename_day_CP1_tosend)

Delay(0,150,mSec)
'-----
'Synchronise clock between datalogger in the packbus

PakBusClock (1)

EndIf
NextScan
EndProg
```

---

## Programme de la centrale météo

```
*****
' CR1000
' MANIP ABIOGEN
' PROGRAMME METEO CONGO + MASTER FILE MGMT
'
' Regis.Burlett@u-bordeaux1.fr
' jerome.bartholome@pierroton.inra.fr
*****
```

'Declare Variables and Units

```
Public outstat_met As Boolean
Public LastFileName_met As String *50

Public time_array(9)
Public YYYY_doy_senddata As String * 20
Public year_yesterday
Public doy_yesterday

Public BattV_met
Public TRHData(2)
Public PAR_in
Public Rain_mm
Public WS_ms
Public WindDir
Public SlrkW
Public SlrMJ

Alias TRHData(1)=AirTC_amb
Alias TRHData(2)=RH_amb

Units BattV_met=Volts
Units PAR_in=umol/m^2/s
```



```
Units Rain_mm=mm
Units WS_ms=meters/second
Units WindDir=Degrees
Units SlrkW=kW/m^2
Units SlrMJ=MJ/m^2
Units AirTC_amb=Deg C
Units RH_amb=%
```

```
.....
```

```
'Define Data Tables
```

```
DataTable(Data_congomet,True,-1)
DataInterval(0,5,Min,10)
```

```
TableFile ("CRD:datacongomet_",9,-1,0,1440,Min,outstat_met,LastFileName_met)
```

```
Minimum(1,BattV_met,FP2,False,False)
Average(1,AirTC_amb,FP2,False)
StdDev(1,AirTC_amb,FP2,False)
Sample(1,RH_amb,FP2)
Average(1,PAR_in,FP2,False)
StdDev(1,PAR_in,FP2,False)
Totalize(1,Rain_mm,FP2,False)
Average(1,WS_ms,FP2,False)
StdDev(1,WS_ms,FP2,False)
Sample(1,WindDir,FP2)
ETsz(AirTC_amb,RH_amb,WS_ms,SlrMJ,12,-5,0,3.2,0,FP2,False)
FieldNames("ETos,Rso")
EndTable
```

```
'Main Program
```

```
BeginProg
Scan(1,Min,3,0)
```

```
'Default Datalogger Battery Voltage measurement BattV
Battery(BattV_met)
```

```
'CS215 Temperature & Relative Humidity Sensor measurements AirTC_amb and RH_amb
SDI12Recorder(TRHData(),7,"0","M!",1,0)
```

```
'ARG100 Tipping Bucket Rain Gauge measurement Rain_mm
PulseCount(Rain_mm,1,1,2,0,0.201,0)
```

```
'03002-5 Wind Speed & Direction Sensor measurements WS_ms and WindDir
PulseCount(WS_ms,1,2,1,1,0.75,0.2)
If WS_ms<0.21 Then WS_ms=0
BrHalf(WindDir,1,mV2500,3,1,1,2500,True,0,_50Hz,355,0)
If WindDir>=360 Then WindDir=0
'CS300 Pyranometer measurements SlrMJ and SlrkW
VoltSe(SlrkW,1,mV250,4,1,0,_50Hz,1,0)
If SlrkW<0 Then SlrkW=0
SlrMJ=SlrkW*0.0015
SlrkW=SlrkW*0.005
'Call Data Tables and Store Data
CallTable(Data_congomet)
```





```
'-----  
' THIS SET OF INSTRUCTION MANAGE THE CLOCK SYNCHRONISATION  
  
If TimeIntoInterval( 450,1440,Min) Then  
    NetworkTimeProtocol ("64.90.182.55 ",0,500)  
EndIf  
  
If TimeIntoInterval( 500,1440,Min) Then  
    ClockReport (ComSDC7,0,2)  
EndIf  
  
If TimeIntoInterval( 505,1440,Min) Then  
    ClockReport (ComSDC7,0,3)  
EndIf  
  
If TimeIntoInterval( 510,1440,Min) Then  
    ClockReport (ComSDC7,0,4)  
EndIf  
  
If TimeIntoInterval( 515,1440,Min) Then  
    ClockReport (ComSDC7,0,5)  
EndIf  
  
NextScan  
SlowSequence  
Scan (1,Hr,3,0)  
  
Delay(1,720,Min)  
RealTime (time_array)  
    If outstat_met = true Then  
  
        'Isolate only the date froml the timestamp in YYYY-MM-DD  
        If time_array(9) >1 Then  
            doy_yesterday = time_array(9) - 1  
            year_yesterday = time_array(1)  
        Else  
            doy_yesterday = 365  
            year_yesterday = time_array(1) - 1  
        EndIf  
  
        YYYY_doy_senddata = year_yesterday + "_" + doy_yesterday  
  
        FileRename (LastFileName_met,Filename_day_met_tosend)  
    '-----  
EndIf  
NextScan  
EndProg
```



## Annexe 4 : Protocole d'extraction d'ADN optimisé en plaque 96 avec le Kit Ademtech.

### **Préparation des Washing Buffer I & II :**

Ajouter dans chaque flacon le volume d'éthanol indiqué.

### **Réactifs supplémentaires nécessaires :**

Isopropanol (froid -20°C), Ethanol pour la préparation des washing Buffer I, II, III, IV

Rnase A (100mg/mL), TE (Tris 10mM – EDTA 1mM – pH8) 0.1X-1X

### **Protocole pour plaque 96 :**

#### **1. Lysis**

- Ajouter 500µL de Lysis Buffer à 5 punchs de feuilles réduits en poudre au Genogrinder.
- Ajouter 4µL de RNase à (100mg/mL). Homogénéiser 30 s, à 1000 rpm au Genogrinder (vérification du fond des tubes).
- Incuber à 56°C pendant 30 minutes sous agitation, puis mettre 5 min à -20°C.
- Ajouter 250 µL de Dichlorométhane, Homogénéiser par pipetage.
- Centrifuger 20 min à 5600g.
- Récupérer 250µL de lysat clarifié et le transférer dans la plaque 96 spécifique.

#### **2. Binding**

- Ajouter 15µL de Silica-MasterBeads puis 250µL d'isopropanol(froid -20°C). Homogénéiser par pipetage.
- Incuber sous agitation 1000 rpm à température ambiante pendant 5 min (Penser au bouchon).

#### **3. Washing I**

- Aimer la suspension de particules pendant au moins 5 min.
- Eliminer le surnageant sans perturber le culot de particules (par pipetage).
- Resuspendre le culot de particules dans 500 µL de Washing Buffer I (avec éthanol).

#### **4. Washing II**

- Aimer la suspension de particules pendant au moins 5 min.
- Eliminer le surnageant sans perturber le culot de particules (par pipetage).
- Resuspendre le culot de particules dans 500 µL de Washing Buffer II (avec éthanol).

#### **5. Washing III**

- Aimer la suspension de particules pendant au moins 5 min.
- Eliminer le surnageant sans perturber le culot de particules (par pipetage).
- Resuspendre le culot de particules dans 500 µL d'éthanol à 70%.

#### **6. Washing IV**

- Aimer la suspension de particules pendant au moins 5 min.
- Eliminer le surnageant sans perturber le culot de particules (par pipetage).
- Resuspendre le culot de particules dans 500 µL d'éthanol à 90%.

#### **7. Séchage**

- Aimer la suspension de particules pendant au moins 5 min.
- Eliminer le surnageant sans perturber le culot de particules.
- Laisser bien sécher à l'air libre pendant 5 min.

#### **8. Elution**

- Reprendre le culot de particules séchées dans 100 µL de TE 0,1X-1X.
- Incuber la suspension sous agitation 1000 rpm à 50°C pendant 5 min.
- Aimer la suspension de particules au moins 5 min.
- Récupérer l'éluât et le transférer dans une plaque PCR.



## Résumé : Déterminisme génétique de la dynamique de croissance et de la composition isotopique du carbone chez l'*Eucalyptus* en réponse aux variations environnementales.

Les différents scénarios sur l'évolution du climat prévoient une augmentation de la fréquence et de l'intensité des sécheresses. La croissance des arbres forestiers étant fortement conditionnée par la disponibilité en eau, ces changements devraient impacter de manière significative la productivité des forêts plantées. La compréhension de l'impact des facteurs génétiques et environnementaux sur la dynamique de croissance est donc un enjeu majeur pour assurer les niveaux de production des plantations de demain. L'*Eucalyptus*, grâce à sa croissance rapide et à la disponibilité de ressources génétiques et génomiques, est un modèle biologique idéal pour mener ces recherches.

L'objectif de cette thèse est de caractériser l'architecture génétique de la dynamique de croissance à différentes échelles de temps chez l'eucalyptus en relation avec : (i) les variations environnementales, et notamment l'évolution de la disponibilité en eau, et (ii) la composition isotopique du carbone de l'arbre ( $\delta^{13}\text{C}$ ), un caractère lié à l'efficacité d'utilisation de l'eau. Pour répondre à cet objectif, un croisement interspécifique *Eucalyptus urophylla* x *E. grandis* a été étudié dans quatre dispositifs expérimentaux en République du Congo. Notre approche se base sur la cartographie des loci à effet quantitatif (QTL) et combine : (i) un génotypage haut débit, (ii) une caractérisation inter et intra-annuelle de la croissance et du  $\delta^{13}\text{C}$ , ainsi qu'un suivi en continu des micro-variations du rayon et (iii) une caractérisation en continu des facteurs environnementaux.

Ces travaux ont tout d'abord conduit à la construction des premières cartes génétiques à haute résolution chez l'*Eucalyptus*. L'analyse de l'architecture génétique du  $\delta^{13}\text{C}$  a ensuite permis de mettre en évidence des gènes candidats positionnels, potentiellement impliqués dans la variation de ce caractère. Enfin, la caractérisation inter et intra-annuelle de la dynamique de croissance a permis de montrer que l'architecture génétique de la croissance, au stade adulte, est structurée par les réponses à l'environnement au stade juvénile. Ces réponses ont ensuite été analysées grâce aux profils de micro-variations du rayon, permettant ainsi de préciser leurs déterminants génétiques. Nos résultats soulignent l'importance de considérer la croissance comme un caractère dynamique, non seulement pour la compréhension de ses bases génétiques, mais également à des fins de sélection de variétés adaptées à un environnement changeant.

**Mots clés :** dynamique de croissance, micro-dendromètre,  $\delta^{13}\text{C}$ , cartographie génétique, QTL, interaction génotype x environnement, *Eucalyptus*.

---

## **Abstract:** Genetic determinism of growth dynamics and carbon isotope composition in *Eucalyptus* in response to environmental changes.

Scenarios of climate changes forecast an increase in frequency and intensity of droughts, related to an increase of global temperatures and changes in rainfall distribution. Growth of forest trees highly depends on water availability and will be significantly impacted by these changes. The understanding of the impact of genetic and environmental factors on the growth dynamics is a major challenge to ensure production levels of future planted forests. *Eucalyptus*, thanks to its rapid growth and the availability of genetic and genomic resources, is a perfect model to conduct this research.

The objective of this thesis is to characterize the genetic architecture of growth dynamics in *Eucalyptus* at different time scales, in relation with: (i) environmental changes, including changes in water availability, and (ii) isotopic composition of carbon ( $\delta^{13}\text{C}$ ), a character associated with water-use efficiency. To this end, an interspecific cross between *E. urophylla* x *E. grandis* was studied in four experimental trials in the Republic of Congo. Our approach, based on mapping of quantitative trait loci (QTL), combines (i) a high-throughput genotyping, (ii) a characterization of inter and intra-annual growth dynamics and  $\delta^{13}\text{C}$ , as well as a continual measurement of stem radial micro-variations and (iii) a continual characterization of environmental factors.

First of all, this work led to the construction of the first high-resolution genetic maps in *Eucalyptus*, improving the sequence of the reference genome. Then, the analysis of genetic architecture of  $\delta^{13}\text{C}$  enabled the identification of positional candidate genes, which might be involved in the variation of this trait. Finally, inter and intra-annual characterization of growth dynamics highlight that genetic architecture of adult growth is structured by responses to the environment at the juvenile stage. These responses were then analyzed using daily profiles of stem radial micro-variations, which enabled the characterization of the genetic determinants of response to the environmental factors at the juvenile stage. Our results highlight the importance of considering growth as a dynamic trait, not only to understand its genetic basis, but also to select in a changing environment.

**Keywords:** growth dynamics, micro-dendrometer,  $\delta^{13}\text{C}$ , genetic mapping, QTL, genotype by environment interaction, *Eucalyptus*